EXTENDED ARRAY EVALUATION PROGRAM. SPECIAL REPORT NO. 2. SIMULATED ON-LINE ADAPTIVE PROCESSING RESULTS USING ALASKA LONG PERIOD ARRAY DATA

Thomas E. Barnard

Texas Instruments, Incorporated

Prepared for:

Air Force Technical Applications Center
Advanced Research Projects Agency

23 October 1973

SIMULATED ON-LINE ADAPTIVE PROCESSING RESULTS USING ALASKA LONG PERIOD ARRAY DATA

SPECIAL REPORT NO. 2

EXTENDED ARRAY EVALUATION PROGRAM

Prepared by
Thomas E. Barnard

TEXAS INSTRUMENTS INCORPORATED
Equipment Group
Post Office Box 6015
Dallas, Texas 75222

Prepared for

AIR FORCE TECHNICAL APPLICATIONS CENTER
AFTAC Project No. VELA T/2705/B/ASD
Alexandria, Virginia 22314

Sponsored by

ADVANCED RESEARCH PROJECTS AGENCY
Nuclear Monitoring Research Office
ARPA Program Code No. 2F10
ARPA Order No. 1714

23 October 1973

Acknowledgment: This research was supported by the Advanced Research Projects Agency, Nuclear Monitoring Research Office, under Project VELA-UNIFORM, and accomplished under the technical direction of the Air Force Technical Applications Center under Contract No. F33657-72-C-0725.

ACKNOWLEDGMENTS

Several people have contributed in various ways to the generation of this report.

Virtually all of the text was typed by Mrs. Cherylann Saunders. Her patience with numerous changes in complicated mathematical equations is sincerely appreciated.

Mr. Charles O. Kemper wrote the computer programs which created Figures III-18 through III-21, III-25 and III-26.

Mr. Terence W. Harley, Captain John W. Woods, and Lieutenant Michael J. Marcus provided encouragement and technical discussions which improved the quality of the report.

Dr. Carl F. Romney asked some pertinent questions which ultimately led to a technique for measuring signal degradation by the adaptive beamformer.

Thomas E. Barnard

ABSTRACT

This report deals with results obtained from operating an adaptive time-domain maximum-likelihood filtering system on data from the Alaska Long-Period Array (ALPA). Signal-to-noise gain of adaptive filtering relative to beamsteering is investigated as a function of convergence rate and steer direction. In addition, the effect upon signal-to-noise gain of freezing the adaptive filter set is described. Both on-azimuth and off-azimuth signals are examined to determine how much they are attenuated in the adaptive-filter beam in comparison with the beamsteer output. Signal-to-noise gain values presented are measured using beam output traces formed from actual seismic data. Theoretical studies of the effect of floating DC levels and roundoff error are also contained in this report.

A major advantage of adaptive multichannel filtering is found to be the capability to narrow the main lobe of the array beam pattern.

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

TABLE OF CONTENTS
(continued)

LIST OF FIGURES

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF TABLES

LIST OF TABLES
(continued)

LIST OF TABLES
(continued)

SECTION I
INTRODUCTION

A. PURPOSE OF THIS STUDY

The adaptive processing task of the Extended Array Evaluation Program has as its objectives:

- To gain experience in operating a real-time adaptive signal estimation processor based on the time-domain maximum-likelihood algorithm
- To perform theoretical studies relating to the convergence of the algorithm and to analyze the output of the adaptive processor in an attempt to upgrade its performance

This report deals solely with the problems associated with operating the real-time adaptive processor on ALPA data. A modified version of the TI interim ALPA system was used to implement the adaptive-filtering algorithm. Theoretical studies of the effect of floating means and roundoff error upon filter performance are contained in this report.

B. DESCRIPTION OF ALASKA LONG-PERIOD ARRAY (ALPA)

The ALPA array is shown in Figure I-1. ALPA is a 19-element hexagonal array with 20 km spacing between sites. Table I-1 gives the ALPA site locations. In this report, sites are referred to in terms of their transmission order. Thus, site 10 refers to site 3-45 in the official nomenclature.

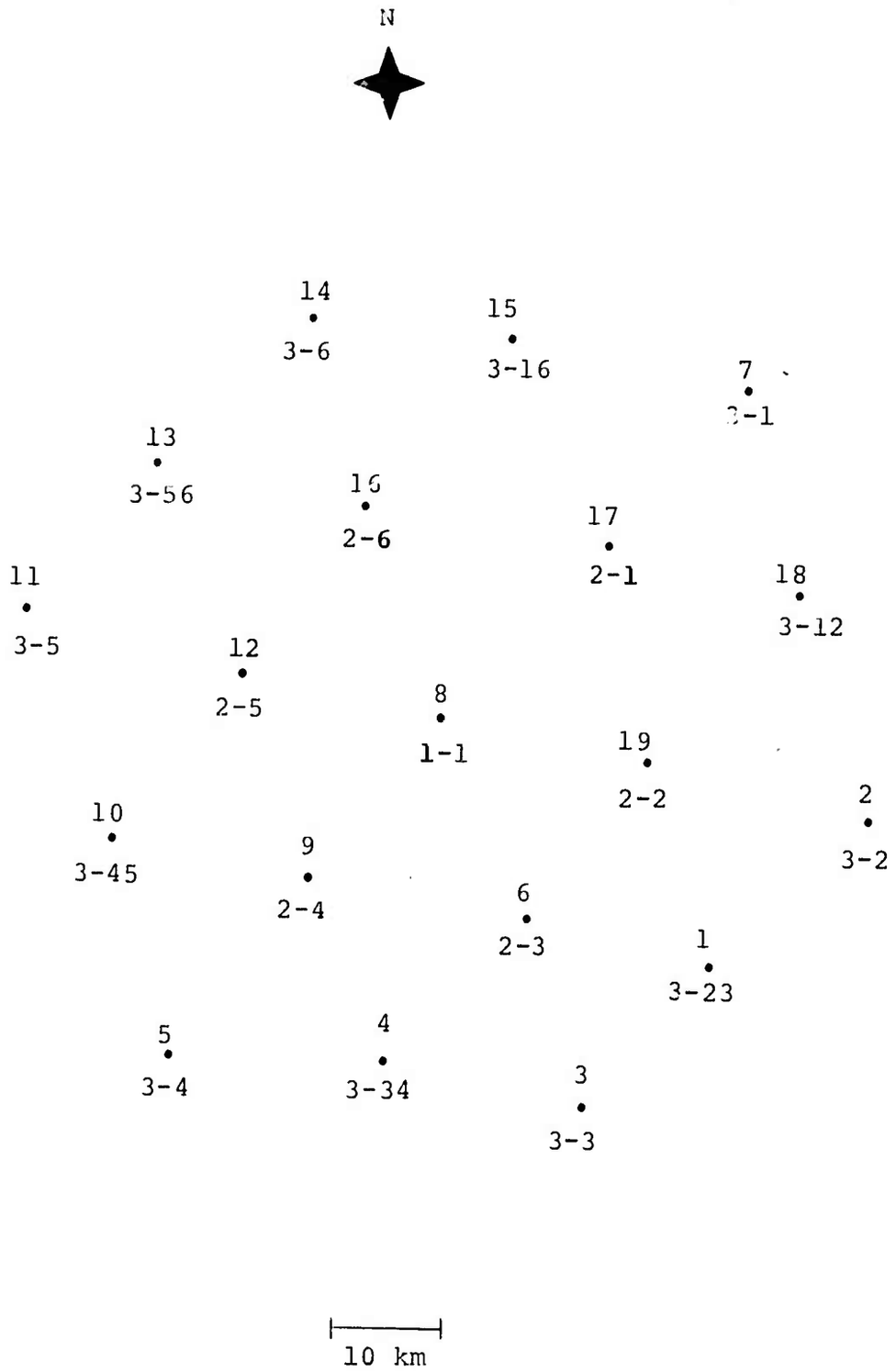


FIGURE I-1
ALPA GEOMETRY

TABLE I-1

ALPA SITE LOCATIONS

Site No.	Transmission Order	N. Lat. ±10 Sec.	W. Long. ±20 Sec.	X Kilometers	Y Kilometers
1-1	8	65°14'00"	147°44'36"	0.0	0.0
2-1	17	65°22'25"	147°24'04"	15.941	15.598
2-2	19	65°11'40"	147°18'58"	19.901	- 4.324
2-3	6	65°03'55"	147°33'50"	8.359	-18.687
2-4	9	65°05'52"	148°00'05"	-12.021	-15.073
2-5	12	65°16'01"	148°08'11"	-18.309	3.737
2-6	16	65°24'02"	147°53'57"	- 7.259	18.594
3-1	7	65°30'10"	147°07'03"	29.153	29.961
3-12	18	65°20'02"	147°00'27"	34.277	11.181
3-2	2	65°08'53"	146°52'45"	40.255	- 9.482
3-23	1	65°01'43"	147°11'46"	25.491	-22.764
3-3	3	64°54'36"	147°26'47"	13.832	-35.953
3-34	4	64°56'41"	147°51'34"	- 5.409	-32.092
3-4	5	64°57'07"	148°17'03"	-25.193	-31.289
3-45	10	65°07'42"	148°24'05"	-30.654	-11.675
3-5	11	65°18'55"	148°35'10"	-39.259	9.112
3-56	13	65°26'09"	148°18'56"	-26.655	22.517
3-6	14	65°33'24"	148°00'00"	-11.956	35.953
3-16	15	65°32'23"	147°35'31"	7.052	34.069

C. ADAPTIVE MULTICHANNEL FILTERING

Multichannel filtering is a form of array processing in which multiple channel inputs undergo individual frequency-shaping and phase-shift filtering prior to the channel-summation operation which produces the beamformer output. Figure I-2 is a schematic diagram of multichannel filtering. This illustration incorporates the option to preprocess the transducer outputs before they are input to the multichannel beamformer. Examples of preprocessing are frequency filtering (most commonly with identical frequency responses on all channels) and time shifting to align waves emanating from a particular direction. The preprocessed transducer outputs become the input channels to a multichannel filter set, where individual filters (generally different from channel to channel) are applied to the input channels. These filters are implemented as convolution filters in time-domain processing or as complex-valued multiplicative filters in frequency-domain processing. The multichannel filter output is created by summing the individual filtered channel outputs.

In systems where second-order statistics (crosscorrelation functions and crosspower spectra) are used to describe interrelationships among the input channels, there are two basic forms of multichannel filtering. In Wiener-Kolmogorov multichannel filtering, the average squared error between the desired signal and the multichannel filter output is minimized. To minimize the mean square error, the crosscorrelation functions or crosspower spectra between the input channels and the desired signal are required. In maximum-likelihood multichannel filtering, the average squared output from the multichannel filter set is minimized subject to signal-preservation constraints which place some suitably-chosen frequency response on the signal. For maximum-likelihood multichannel filtering, unlike Wiener-Kolmogorov filtering, only the direction of the signal needs to be specified, but not the signal-to-noise ratio.

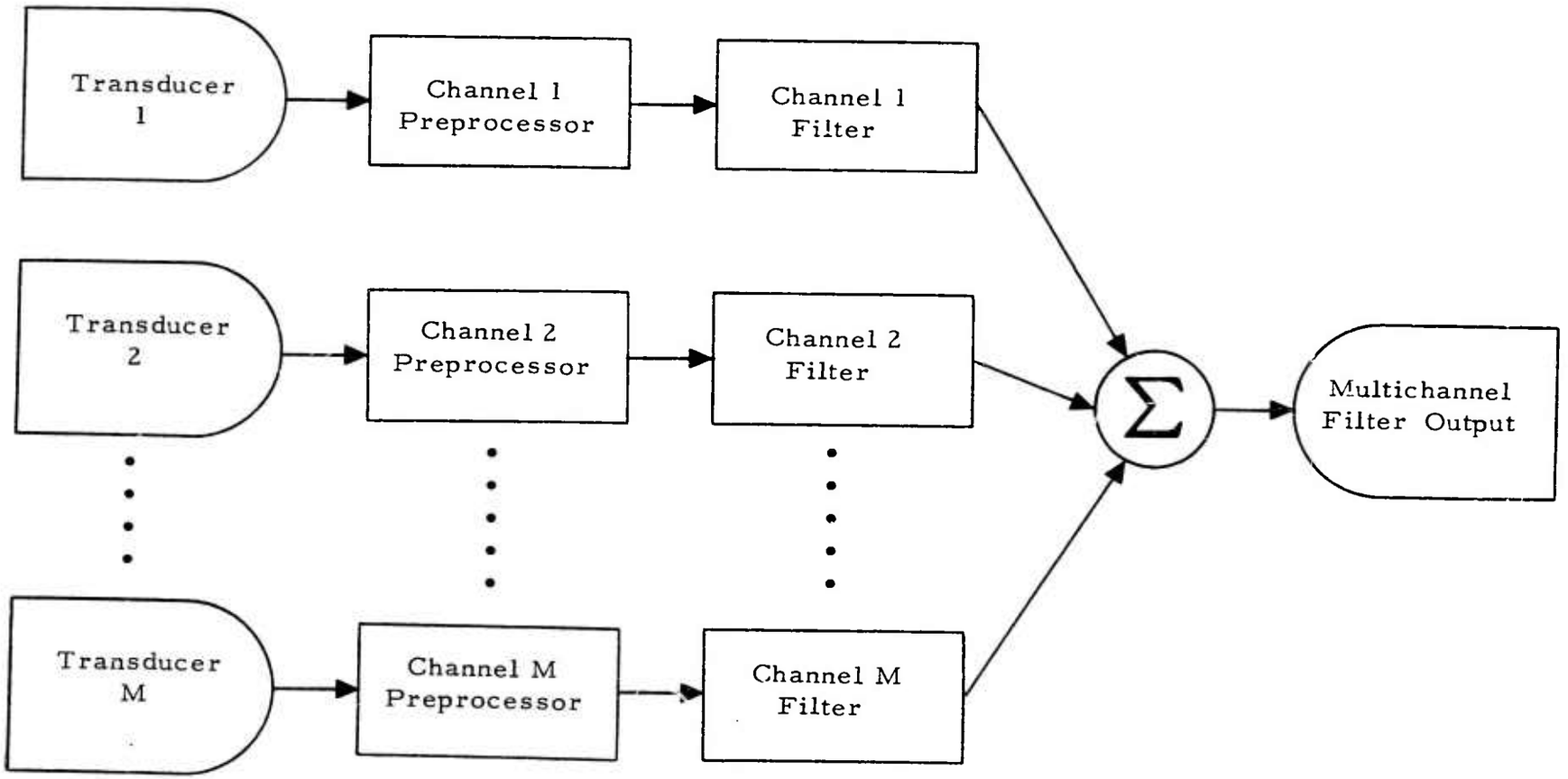


FIGURE I-2

SCHEMATIC DIAGRAM OF MULTICHANNEL FILTERING

Multichannel filtering can be employed with fixed or time-varying filter sets. When the filters are updated as new data inputs enter the multichannel processor, the process is called adaptive filtering. Adaptive-filtering algorithms with significant computational advantages over fixed multichannel filtering are available. When the inputs to the multichannel processor are time-stationary (in the wide sense), these algorithms yield filter sets which converge in the mean to the corresponding fixed multichannel filter sets. After adaptive filter sets reach the vicinity of the corresponding fixed filter sets, they fluctuate about the fixed-filter solution in the presence of time-stationary data: the adaptive filters converge in the mean in the sense that the average position of the fluctuating adaptive filters is identical to the fixed-filter solution. When the statistics of the data entering the multichannel processor slowly change with time, adaptive filtering can react to the changes in a semi-continuous manner. If fixed filtering is used in this situation, newly-designed filters change in a more abrupt fashion. When, as in this case, the statistics of the data shift with time, the adaptive-filter solution lags behind the fixed-filter solution corresponding to the instantaneous statistics. The extent of the lag can be controlled by changing the adaptation rate. The choice of an adaptation rate involves a tradeoff between misadjustment (higher-than-optimum error or power due to the adaptive-filter fluctuations) and the lag behind the optimum instantaneous fixed-filter solution. A different kind of lag occurs when fixed filter sets are periodically redesigned: statistics must be accumulated over a design interval so that, as a result, the fixed-filter solution cannot be implemented until the next design interval.

In the conventional technique of array processing, simple time delays or phase shifts are applied to the input channels before summing to generate the beam output. Optimum multichannel filtering introduces considerable new flexibility into the beamforming process. Since it is possible to weight

the input channels differently, channels with higher signal-to-noise ratios can be emphasized at the expense of noisier channels. When well instrumented arrays are utilized, this capability is generally of minor importance. A far more consequential feature of adaptive filtering is the ability to form array antenna patterns which optimally pass a signal while simultaneously rejecting propagating noise. Deep nulls can be aimed toward off-beam noise sources. When strong off-azimuth noise sources are present, the creation of such nulls is an automatic result of the optimality of the multichannel processor. The conventional time-shift-and-sum or phase-shift processor, in contrast, has a beam pattern determined solely by the steer direction and the array geometry.

In most cases, the potential improvement of optimum multichannel filtering relative to beamsteering is determined by the coherence of the noise field across an array: the greater the similarity of the noise field from channel to channel, the greater is the optimum-multichannel-filter improvement over beamsteering. When, on the other hand, noise is completely uncorrelated between sensors and identical signal and noise power levels are encountered at all array sites, there is no potential for improvement: in this case, the optimum filter set is a beamsteer processor. The decision to employ or not to employ an optimum-filter technique of processing depends critically on measurements of the noise field at any given array. Once these measurements are available, the additional cost of implementing an optimum-filter system can be quantitatively weighed against the advantages of greater noise suppression relative to the conventional beamsteer processing technique.

D. DESCRIPTION OF THE MAXIMUM-LIKELIHOOD ADAPTIVE FILTER ALGORITHM

The adaptive-filter output $y(t)$ at time t is formed by applying a convolution filter to each channel and summing the outputs of all channels:

$$y(t) = \sum_{i=1}^{M} \sum_{j=-N}^{N} a_i(j) x_i(t-j)$$

where $a_i(j)$ is the filter weight for the i-th channel at a lag of j sample points, $x_i(t-j)$ is the value of the channel i at time t-j. M is the number of channels, and 2N+1 is the total length of the filter in points. Prior to forming the filter output, each channel is time-shifted to time-align energy arriving from the desired steer direction.

The adaptive filter weights are updated by the following algorithm:

$$a_i^{new}(j) = a_i^{old}(j) + \lambda(t)\, y(t) \left[\bar{x}(t-j) - x_i(t-j) \right]$$

where

$$\bar{x}(t-j) = \frac{1}{M} \sum_{i=1}^{M} x_i(t-j)$$

and $\lambda(t)$ is the convergence parameter at time t. This update algorithm incorporates the maximum-likelihood constraints.

The convergence parameter $\lambda(t)$ is calculated by the formula

$$\lambda(t) = \frac{2K_s}{(2N+1) \sum_{i=1}^{M} P_i(t)}$$

where K_s is an input parameter, and $P_i(t)$ is a moving power average for the i-th channel. $P_i(t)$ is computed by the formula:

$$P_i(t) = (1 - \mu) \left[\bar{x}(t) - x_i(t) \right]^2 + \mu P_i(t-1) \qquad t \geq 1$$

where μ is an input parameter. $P_i(0)$ is zero, and several values of $P_i(t)$ are computed before the filter is allowed to vary.

E. GEOMETRICAL INTERPRETATION OF THE ABF ALGORITHM

In vector form, the adaptive-beamforming filter update equation may be written

$$A^{new} - A^{old} = \lambda(t)\, X^T A^{old} (\bar{X} - X) ,$$

where the superscript T denotes transposition, and where the filter weight vector A, the data vector X, and the beamsteer output vector $\bar{X}$ are, respectively,

$$A = \begin{bmatrix} \begin{bmatrix} a_1(-N) \\ \vdots \\ a_M(-N) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_1(0) \\ \vdots \\ a_M(0) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_1(N) \\ \vdots \\ a_M(N) \end{bmatrix} \end{bmatrix}, \quad X = \begin{bmatrix} \begin{bmatrix} x_1(t+N) \\ \vdots \\ x_M(t+N) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} x_1(t) \\ \vdots \\ x_M(t) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} x_1(t-N) \\ \vdots \\ x_M(t-N) \end{bmatrix} \end{bmatrix}, \text{ and } \bar{X} = \begin{bmatrix} \begin{bmatrix} \bar{x}(t+N) \\ \vdots \\ \bar{x}(t+N) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \bar{x}(t) \\ \vdots \\ \bar{x}(t) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \bar{x}(t-N) \\ \vdots \\ \bar{x}(t-N) \end{bmatrix} \end{bmatrix}.$$

The objective of maximum-likelihood adaptive beamforming is to reduce the average squared filter output

$$\overline{y^2(t)} = \overline{(A^T X)(X^T A)} = A^T \overline{XX^T} A$$

subject to a set of signal-preservation constraints on the filter vector A. After preshifting the input channels to time-align energy from the look direction, these constraints can be written

$$\sum_{i=1}^{M} a_i(j) = d(j) \qquad (j = -N, \ldots, -1, 0, 1, \ldots, N),$$

where the constants $d(j)$ specify a convolution filter having the desired frequency response on a signal from the steer direction. When such a signal $s(t)$ appears in identical form on all channels, the signal output from the beamformer is

$$\sum_{j=-N}^{N} \left[\sum_{i=1}^{M} a_i(j) \right] s(t-j) = \sum_{j=-N}^{N} d(j)s(t-j) .$$

For the adaptive beamforming employed in this study, a white frequency response is specified by setting

$$d(j) = \delta_{jo} \qquad (j = -N, \ldots, -1, 0, 1, \ldots, N),$$

where δ_{jo} is the Kronecker delta operator

$$\delta_{jo} = \begin{cases} 1 & \text{if } j = 0 \\ 0 & \text{if } j \neq 0. \end{cases}$$

To reduce the average squared filter output $\overline{y^2(t)}$, the method of steepest descent (with two modifications) is used. In the unmodified form of the method of steepest descent, the filter vector A moves in the direction opposite to the gradient of the average squared filter output $\overline{y^2(t)}$:

$$A^{new} - A^{old} = -c\nabla(A^T \overline{XX^T} A) = -2c\,\overline{XX^T}\,A .$$

The first modification is to replace the crosscorrelation matrix $\overline{XX^T}$ with the rank-one matrix XX^T formed from the instantaneous vector X at time t:

$$A^{new} - A^{old} = -2c\,X(X^T A) = -2c\,X\,y(t) .$$

This modification of Widrow (Widrow, 1966) approximates, in effect, the crosscorrelation-matrix time averaging through successive applications of the filter update algorithm. The approximation becomes increasingly accurate as the rate of change of the filter vector A is slowed by reducing the scalar convergence parameter c. The filter vector A converges in the mean under suitably prescribed conditions (Daniell, 1968) to the vector obtained by using the crosscorrelation matrix $\overline{XX^T}$ in the update equation. Ultimately the filter vector A oscillates about its mean. The size of the oscillations can be controlled by varying the parameter c. The reason for the Widrow modification is a reduction in the computational operations required for the filter update from a number of proportional to the square of the dimension of the vectors X and A to a number linearly proportional to their dimension.

The second modification of the steepest-descent method is to alter the direction of the vector $-c\nabla\left[y^2(t)\right]$ so that the ensuing update vector $(A^{new} - A^{old})$ is the vector nearest to $-c\nabla\left[y^2(t)\right]$ which satisfies the constraints on the filter update vector. Since the sum across channels

$$\sum_{i=1}^{M} a_i(j)$$

of the filter vector A is a fixed value $d(j)$ at any lag value j, the filter update vector $(A^{new} - A^{old})$ must sum to zero at each lag:

$$\sum_{i=1}^{M}\left[a_i^{new}(j) - a_i^{old}(j)\right] = 0 \quad (j = -N, \ldots, -1, 0, 1, \ldots, N).$$

The filter update vector must be perpendicular to each of the $(2N+1)$ unit vectors U_k specified by their components

$$\left[u_i(j)\right]_k = \delta_{jk}/\sqrt{M} \ ,$$

where δ_{jk} is the Kronecker delta operator. Each of the vectors U_k is zero except in the k-th lag position, where all components are equal to $1/\sqrt{M}$. In vector form, the filter update vector must satisfy the (2N+1) constraints

$$U_k^T (A^{new} - A^{old}) = 0 \qquad (j = -N, \ldots, -1, \ 0, \ 1, \ldots, N).$$

The vector $\bar{X}$, since it has identical components at any lag value j, is a linear combination of the vectors U_k and is perpendicular to all possible update vectors satisfying the constraint conditions. The vector $(X-\bar{X})$, on the other hand, satisfies the constraint conditions:

$$U_k^T(X-\bar{X}) = \sum_{i=1}^{M} \sum_{j=-N}^{N} \delta_{jk} \left[x_i(j) - \bar{x}(j)\right]$$

$$= \sum_{i=1}^{M} \left[x_i(k) - \bar{x}(k)\right] = \left[\sum_{i=1}^{M} x_i(k)\right] - M\bar{x}(k)$$

$$= M\bar{x}(k) - M\bar{x}(k) = 0 \ .$$

Thus the vector X can be resolved into two mutually orthogonal components $\bar{X}$ (perpendicular to the constraint space for the filter update vector) and $X-\bar{X}$ (lying within the constraint space). The negative $-c\nabla\left[y^2(t)\right]$ of the scaled gradient of $y^2(t)$ is a scalar multiple $-\left[2cy(t)\right]X$ of the vector X. The nearest point to $-c\nabla\left[y^2(t)\right]$ on the constraint space is the vector $\left[2cy(t)\right](\bar{X}-X)$ formed by subtracting the component $-\left[2cy(t)\right]\bar{X}$ perpendicular to the constraint space from the scaled negative gradient vector $-\left[2cy(t)\right]X$. The final form of the filter update equation is, therefore,

$$A^{new} - A^{old} = \left[2 c y(t)\right] (\bar{X} - X) .$$

This situation is illustrated in Figure I-3. The resultant filter update vector is the projection of $-c \nabla \left[y^2(t)\right]$ onto the constraint space.

F. A SIMPLE ILLUSTRATIVE EXAMPLE

Suppose that three channels are input to an adaptive beamformer with one lag per channel. Signals are defined to be simultaneous spikes on all three channels. Noise, on the other hand, appears in the form of unit-amplitude spikes on the second channel. When signal and noise occur, they are characterized by the respective vectors

$$S = \begin{bmatrix} x_1(t) = s \\ x_2(t) = s \\ x_3(t) = s \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} x_1(t) = 0 \\ x_2(t) = 1 \\ x_3(t) = 0 \end{bmatrix}.$$

Figure I-4 depicts time series inputs for this simple example. The adaptive filter set is initialized with beamsteer weights:

$$A = \begin{bmatrix} a_1(0) = 1/3 \\ a_2(0) = 1/3 \\ a_3(0) = 1/3 \end{bmatrix}.$$

An optimum filter set is

$$A = \begin{bmatrix} a_1(0) = 1/2 \\ a_2(0) = 0 \\ a_3(0) = 1/2 \end{bmatrix}.$$

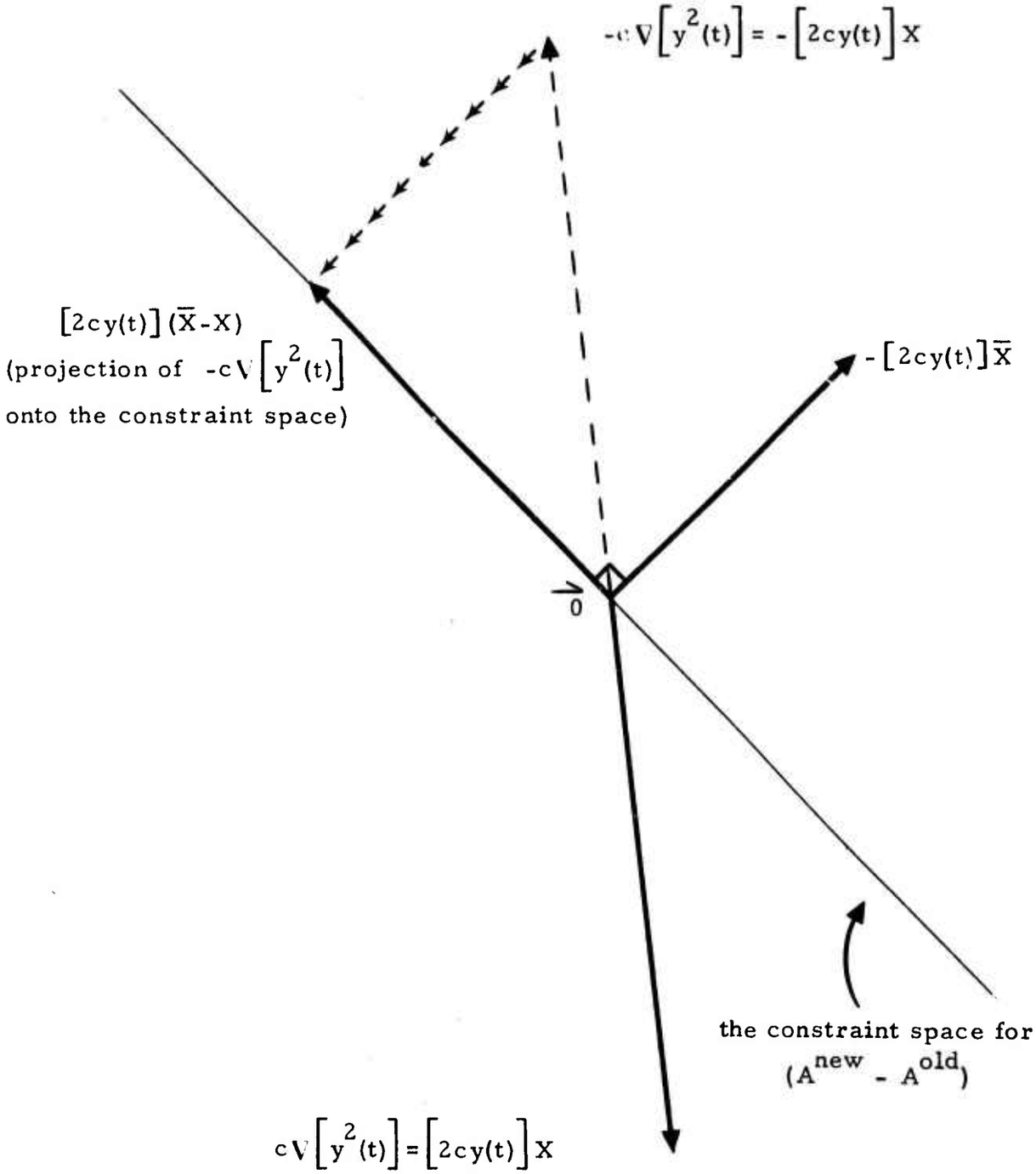
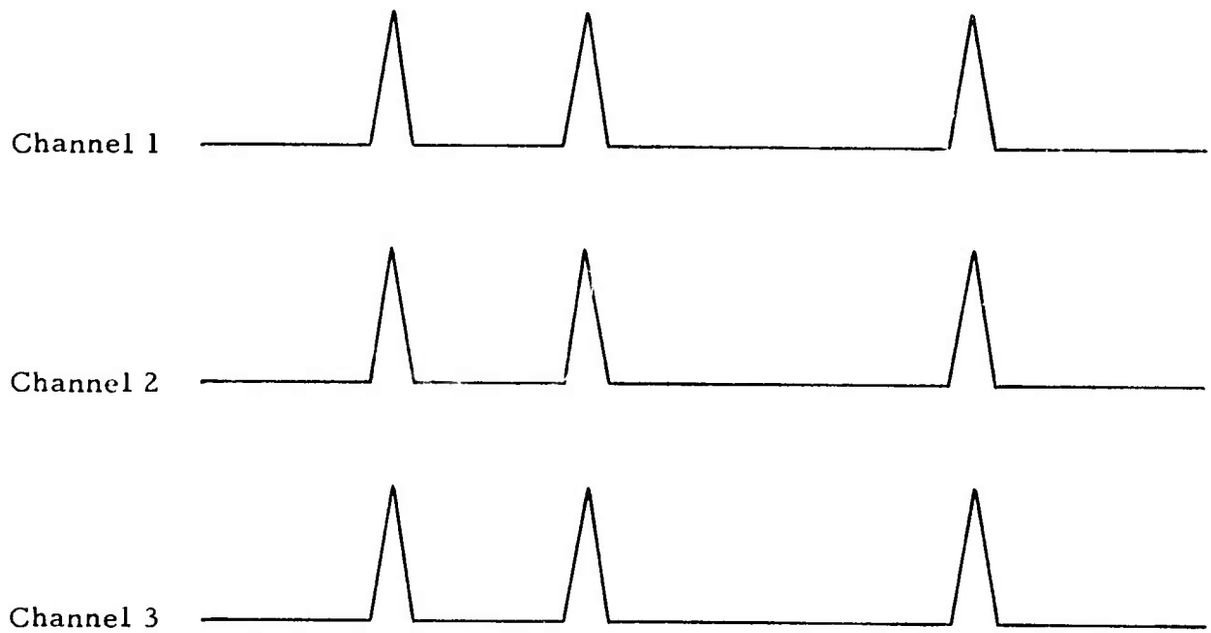
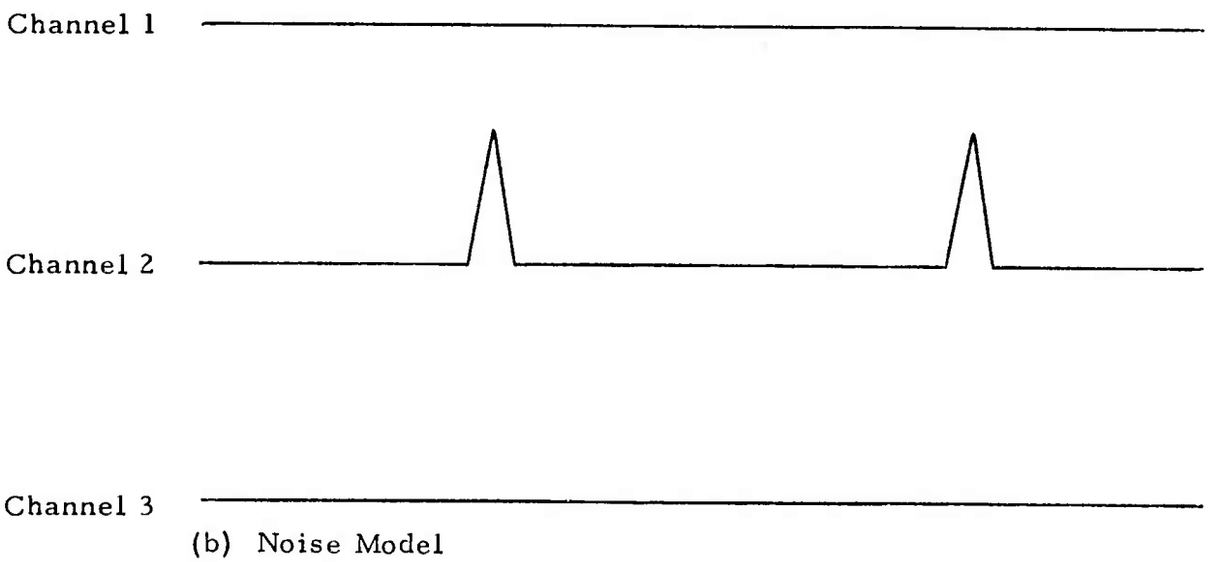


FIGURE I-3

A GEOMETRICAL INTERPRETATION OF TIME-DOMAIN MAXIMUM-LIKELIHOOD ADAPTIVE FILTERING



(a) Signal Model



(b) Noise Model

FIGURE I-4

A SIMULATED DISPLAY OF THE TIME SERIES INPUT CORRESPONDING TO THE ILLUSTRATIVE EXAMPLE

Whenever a noise spike first appears on channel 2, the filter vector A is updated according to the equation

$$A^{new} = A^{old} + 2cy(t)\,(\bar{N}-N)$$

or

$$A^{new} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} + 2c \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{3} - 0 \\ \frac{1}{3} - 1 \\ \frac{1}{3} - 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} + \frac{2}{3}c \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3} + \frac{2}{9}c \\ \frac{1}{3} - \frac{4}{9}c \\ \frac{1}{3} + \frac{2}{9}c \end{bmatrix}.$$

The adaptive update equation shifts its weighting from channel 2 to channels 1 and 3. The second time the noise spike appears on channel 2, the filter output is

$$y(t) = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} + \frac{2}{9}c \\ \frac{1}{3} - \frac{4}{9}c \\ \frac{1}{3} + \frac{2}{9}c \end{bmatrix} = \frac{1}{3} - \frac{4}{9}c .$$

The amplitude of the noise output is diminished whenever $0 < c < 3/2$. After k updates, the adaptive filter vector is

$$\begin{bmatrix} \frac{1}{2} - \frac{1}{6}\left(1 - \frac{4}{3}c\right)^k \\ \frac{1}{3}\left(1 - \frac{4}{3}c\right)^k \\ \frac{1}{2} - \frac{1}{6}\left(1 - \frac{4}{3}c\right)^k \end{bmatrix}.$$

When the (k+1)-st noise spike appears on channel 2, the output of the adaptive beamformer is

$$y(t) = \frac{1}{3}\left(1 - \frac{4}{3}c\right)^k .$$

Each time a new noise spike appears on channel 2, the adaptive filter output is $(1 - 4c/3)$ times its previous value. After each new noise spike, the difference between the optimum filter vector and the old filter vector is reduced by the same factor $(1 - 4c/3)$. In this simple example where the noise crosscorrelation matrix $\overline{NN^T}$ has a single non-zero eigenvalue, the average squared filtered noise output $\overline{y^2(t)}$ is minimized when $c = 3/4$. In this case, noise is completely eliminated starting with the second noise spike on channel 2. When $0 < c < 3/4$, the noise output always has the same sign as the spike on channel 2. When $c > 3/4$, however, the noise output

alternately differs and agrees in sign with the input spike on channel 2. The magnitude of the output spike always diminishes as long as $0 < c < 3/2$. When $c > 3/2$, on the other hand, the adaptive filter vector diverges and the ABF output increases in magnitude with the arrival of each new noise spike.

In this idealized example, the signal does not affect the filter update since

$$\bar{S} - S = \begin{bmatrix} s \\ s \\ s \end{bmatrix} - \begin{bmatrix} s \\ s \\ s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

In the event that a signal spike appears simultaneously on all channels, the output from the beamformer is s since the filter weights sum to one.

G. ORGANIZATION OF REPORT

Section II describes the problem of floating means in ALPA data, its effect on filter performance, and the techniques used to eliminate it. Simple subtraction of the channel means did not adequately solve the problem. This section is intended for data analysts and programmers who must deal with means in the input data.

Section III examines the effect of roundoff error in an integer-arithmetic implementation of the filter update equation. Simple approximations for the average angle of error in the vectors $(\bar{X}-X)$ and $(A^{new}-A^{old})$ are presented. The angle of error in $(\bar{X}-X)$ is affected by the number of bits used to represent the data values $x_i(t-j)$, whereas the angle of error in $(A^{new} - A^{old})$ is also influenced by the number of bits employed to represent the filter weights $a_i(j)$. The simple approximations given for the angular error in $(\bar{X}-X)$ and $(A^{new} - A^{old})$ are probably adequate to specify the data-value and filter-weight representations to the nearest four bits when directional

error limits are placed on the vectors $(\bar{X}-X)$ and $(A^{new} - A^{old})$. This section is intended primarily for digital design engineers and programmers who may need to consider roundoff error when implementing an integer-arithmetic adaptive-filtering system.

In Section IV, the subject of investigation is the ability of adaptive multichannel filtering to provide signal-to-noise ratio improvement relative to beamsteering in the presence of background noise. Signal degradation and noise reduction measurements from real data are combined to yield the associated signal-to-noise gain. The convergence rates where the highest signal-to-noise gains occur are ascertained in this way. This section is intended to assist those persons evaluating the potential usefulness of an adaptive-filtering system for processing ALPA data.

Section V studies the effect of freezing the adaptive filter set. Loss in noise reduction is determined by comparing the adaptive-filter beam output when the filter is allowed to vary with the adaptive-filter beam output when the filter is frozen. Signal degradation is remeasured when the adaptive filter set is frozen in order to estimate the resulting signal-to-noise gain. In processing the data samples used to evaluate the filter-freeze procedure, several off-azimuth events are also run through the adaptive beamformer. The potential improvement in interfering-event situations is demonstrated as a byproduct of the investigations presented in this section. The results of this section are intended both for data analysts and programmers interested in the effect of the filter-freeze procedure and for those persons interested in the effectiveness of the adaptive-filter algorithm in processing interfering events.

Section VI examines noise reduction achieved by adaptive filtering for different look directions and discusses variations in processing gain for the ten four-hour noise samples processed for this report. This section is intended to give some idea of the variability which can be expected in the

performance of the adaptive-filtering process when background noise is present. This information should be useful in evaluating the potential improvements to be gained through adaptive filtering.

Section VII presents the conclusions of this study. For those with insufficient time to analyze in detail the results of the individual sections in this report, this section gives the highlights of this investigation.

SECTION II
THE PROBLEM OF FLOATING MEANS

A. FLOATING MEANS IN ALPA DATA

In data recorded prior to the summer of 1972, some of the digitized traces transmitted from ALPA had mean levels higher than 1000 computer counts. These mean levels were contrasted with RMS levels in tens of computer counts (after mean removal). Furthermore, these DC levels gradually changed over a period of a few hours. Subtraction of a fixed DC level was insufficient to reduce the mean to a level significantly below the RMS level of any given trace because of the gradually changing mean levels. Even the removal of an exponentially smoothed running mean produced similar results because of a time lag between the mean computation and mean removal.

The floating means in ALPA data are caused by drift in the preamplifiers at the individual sites of the array. In the summer of 1972, the original preamplifiers were replaced by new preamplifiers of different design. The effect of this substitution was to reduce the floating mean levels by an order of magnitude: while mean levels with the original equipment could sometimes reach thousands of computer counts, mean levels were reduced to levels expressed in hundreds of counts (typically between 100 and 200 computer counts) with the change in instrumentation.

Although the problem of floating means has been considerably ameliorated by the dramatic reduction in mean levels at ALPA, the current DC levels still lie above the channel RMS amplitude levels after mean removal. The same problem is evident at the NORSAR array. In fact, the floating-mean problem may be a general problem with long-period seismometers.

Consequently, effective techniques for removing the floating DC levels appear to be essential in processing long-period data with adaptive-filtering algorithms.

On a quiet noise sample from day 203 of 1971, negative improvement from adaptive filtering (as measured in terms of noise reduction) was consistently obtained until the problem of floating means was studied and dealt with effectively. The chief symptom of the trouble was a particularly poor performance at frequencies below the frequency equal to the reciprocal of the filter length.

Subsection B presents a theoretical study of the effects of DC bias on maximum-likelihood adaptive filtering. In Subsection B, the results are derived using the fixed-filter maximum-likelihood multichannel-filter design equations. Since the adaptive-filter solution converges in the mean to the fixed-filter solution, similar if not precisely identical results can be expected in the case of adaptive filtering. Empirical verification of this fact can be derived from the elimination of the predicted floating-mean performance symptoms upon removal of the floating means from the data.

Subsection C describes the measures taken to eliminate problems associated with DC bias. A necessary remedy was the application of a prefilter with precisely zero response at 0 Hz. In addition, computational bias in the adaptive-filter computer program had to be reduced to the maximum extent possible.

B. EFFECT OF DC LEVELS ON ADAPTIVE FILTERING IMPROVEMENT

The general design equation for two-channel, (2N+1)-point maximum-likelihood multichannel filter set is

$$
\begin{bmatrix}
\begin{bmatrix} \phi_{11}(0) & \phi_{12}(0) & 1 \\ \phi_{21}(0) & \phi_{22}(0) & 1 \\ 1 & 1 & 0 \end{bmatrix} & \begin{bmatrix} \phi_{11}(-1) & \phi_{12}(-1) & 0 \\ \phi_{21}(-1) & \phi_{22}(-1) & 0 \\ 0 & 0 & 0 \end{bmatrix} & \cdots & \begin{bmatrix} \phi_{11}(-2N) & \phi_{12}(-2N) & 0 \\ \phi_{21}(-2N) & \phi_{22}(-2N) & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
\begin{bmatrix} \phi_{11}(1) & \phi_{12}(1) & 0 \\ \phi_{21}(1) & \phi_{22}(1) & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} \phi_{11}(0) & \phi_{12}(0) & 1 \\ \phi_{21}(0) & \phi_{22}(0) & 1 \\ 1 & 1 & 0 \end{bmatrix} & \cdots & \begin{bmatrix} \phi_{11}(1-2N) & \phi_{12}(1-2N) & 0 \\ \phi_{21}(1-2N) & \phi_{22}(1-2N) & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
\vdots & \vdots & \ddots & \vdots \\
\begin{bmatrix} \phi_{11}(2N) & \phi_{12}(2N) & 0 \\ \phi_{21}(2N) & \phi_{22}(2N) & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} \phi_{11}(2N-1) & \phi_{12}(2N-1) & 0 \\ \phi_{21}(2N-1) & \phi_{22}(2N-1) & 0 \\ 0 & 0 & 0 \end{bmatrix} & \cdots & \begin{bmatrix} \phi_{11}(0) & \phi_{12}(0) & 1 \\ \phi_{21}(0) & \phi_{22}(0) & 1 \\ 1 & 1 & 0 \end{bmatrix}
\end{bmatrix}
\begin{bmatrix} a_1(-N) \\ a_2(-N) \\ -\lambda(-N) \\ \vdots \\ a_1(0) \\ a_2(0) \\ -\lambda(0) \\ \vdots \\ a_1(N) \\ a_2(N) \\ -\lambda(N) \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

where $\phi_{ij}(\tau)$ denotes the crosscorrelation function between channel i and channel j at time lag τ, $a_j(\tau)$ denotes the filter point for channel j at time lag τ, and $\lambda(\tau)$ denotes the Lagrangian multiplier associated with the maximum-likelihood constraint condition imposed upon the filter weights for time lag τ. The Lagrangian multiplier $\lambda(0)$ for zero lag is equal to the mean square error and noise power output of the filter set. A derivation of the design equation is given in part 2 of Subsection III-D.

Assume two traces each consist of a mean m_i with white random noise of power p_i. Then

$$\phi_{ij}(\tau) = m_i m_j \qquad \tau \neq 0$$

$$\phi_{ij}(0) = m_i m_j \qquad i \neq j$$

$$\phi_{ii}(0) = m_i^2 + p_i$$

Let

$$P = \begin{bmatrix} p_1 & 0 & 1 \\ 0 & p_2 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

$$M = \begin{bmatrix} m_1 \\ m_2 \\ 0 \end{bmatrix},$$

$$O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ,$$

$$A(\tau) = \begin{bmatrix} a_1(\tau) \\ a_2(\tau) \\ -\lambda(\tau) \end{bmatrix} ,$$

and

$$\Delta(\tau) = \begin{bmatrix} 0 \\ 0 \\ \delta_{o\tau} \end{bmatrix} ,$$

where δ denotes the Kronecker delta.

The design equation for a symmetrical two-channel, (2N+1)-point maximum-likelihood filter set, for the noise field specified, reduces to

$$\left\{ \begin{bmatrix} P & O & \cdots & O & O \\ O & P & & O & O \\ \vdots & & \ddots & & \vdots \\ O & O & & P & O \\ O & O & \cdots & O & P \end{bmatrix} + \begin{bmatrix} M \\ M \\ \vdots \\ M \\ M \end{bmatrix} \begin{bmatrix} M^T & M^T & \cdots & M^T & M^T \end{bmatrix} \right\} \begin{bmatrix} A(-N) \\ \vdots \\ A(0) \\ \vdots \\ A(N) \end{bmatrix} = \begin{bmatrix} \Delta(-N) \\ \vdots \\ \Delta(0) \\ \vdots \\ \Delta(N) \end{bmatrix}$$

Let Π denote the (2N+1) by (2N+1) block matrix with diagonal element P. Let μ denote the (2N+1) block column vector with element M. Then the total noise matrix is $\Phi = \Pi + \mu\mu^T$.

$$\Phi^{-1} = \Pi^{-1} - \frac{(\Pi^{-1}\mu)(\mu^T\Pi^{-1})}{(1+\mu^T\Pi^{-1}\mu)}$$

$$\Pi^{-1} = \begin{bmatrix} P^{-1} & O & \cdots & O & O \\ O & P^{-1} & & O & O \\ \vdots & & \ddots & & \vdots \\ O & O & & P^{-1} & O \\ O & O & \cdots & O & P^{-1} \end{bmatrix}$$

$$\Pi^{-1}\boldsymbol{\mu} = \Pi^{-1}\begin{bmatrix} M \\ M \\ \vdots \\ M \\ M \end{bmatrix} = \begin{bmatrix} P^{-1}M \\ P^{-1}M \\ \vdots \\ P^{-1}M \\ P^{-1}M \end{bmatrix}$$

$$\boldsymbol{\mu}^T\Pi^{-1}\boldsymbol{\mu} = \begin{bmatrix} M^T & M^T & \cdots & M^T & M^T \end{bmatrix}\begin{bmatrix} P^{-1}M \\ P^{-1}M \\ \vdots \\ P^{-1}M \\ P^{-1}M \end{bmatrix} = (2N+1)M^TP^{-1}M$$

$$\Pi^{-1}\boldsymbol{\mu}\boldsymbol{\mu}^T\Pi^{-1} = \begin{bmatrix} P^{-1}M \\ P^{-1}M \\ \vdots \\ P^{-1}M \\ P^{-1}M \end{bmatrix}\begin{bmatrix} M^TP^{-1} & M^TP^{-1} & \cdots & M^TP^{-1} & M^TP^{-1} \end{bmatrix}$$

$$P^{-1}M = \frac{1}{p_1+p_2}\begin{bmatrix} 1 & -1 & p_2 \\ -1 & 1 & p_1 \\ p_2 & p_1 & -p_1p_2 \end{bmatrix}\begin{bmatrix} m_1 \\ m_2 \\ 0 \end{bmatrix}$$

$$= \frac{1}{p_1+p_2}\begin{bmatrix} m_1 - m_2 \\ m_2 - m_1 \\ p_2m_1 + p_1m_2 \end{bmatrix}$$

$$M^TP^{-1}M = \frac{1}{p_1+p_2}\begin{bmatrix} m_1 & m_2 & 0 \end{bmatrix}\begin{bmatrix} m_1 - m_2 \\ m_2 - m_1 \\ p_2m_1 + p_1m_2 \end{bmatrix} = \frac{(m_1 - m_2)^2}{p_1 + p_2}$$

$$P^{-1}MM^TP^{-1} = \frac{(m_1 - m_2)^2}{(p_1 + p_2)^2}\begin{bmatrix} 1 \\ -1 \\ \boldsymbol{\beta} \end{bmatrix}\begin{bmatrix} 1 & -1 & \boldsymbol{\beta} \end{bmatrix}$$

where $\beta = \dfrac{p_2 m_1 + p_1 m_2}{m_1 - m_2}$.

$$\Phi^{-1} = \Pi^{-1} - \frac{(\Pi^{-1}\mu)(\mu^T \Pi^{-1})}{(1+\mu^T \Pi^{-1}\mu)}$$

$$1 + \mu^T \Pi^{-1} \mu = 1 + \frac{(2N+1)(m_1 - m_2)^2}{p_1 + p_2}$$

$$\Phi^{-1} = \frac{1}{p_1+p_2} \left\{ \begin{bmatrix} (p_1+p_2)P^{-1} & O & \cdots & O & O \\ O & (p_1+p_2)P^{-1} & & O & O \\ \vdots & & \ddots & & \vdots \\ O & O & & (p_1+p_2)P^{-1} & O \\ O & O & \cdots & O & (p_1+p_2)P^{-1} \end{bmatrix} - \frac{(m_1 - m_2)^2}{\left[(p_1 + p_2) + (2N + 1)(m_1 - m_2)^2\right]} \begin{bmatrix} E & E & \cdots & E & E \\ E & E & & E & E \\ \vdots & & \ddots & & \vdots \\ E & E & & E & E \\ E & E & \cdots & E & E \end{bmatrix} \right\}$$

where $$E = \begin{bmatrix} 1 & -1 & \beta \\ -1 & 1 & -\beta \\ \beta & -\beta & \beta^2 \end{bmatrix}$$

$$\Phi^{-1} = \frac{1}{p_1+p_2} \left\{ \begin{bmatrix} D & O & \cdots & O & O \\ O & D & & O & O \\ \vdots & & \ddots & & \vdots \\ O & O & & D & O \\ O & O & \cdots & O & D \end{bmatrix} + \gamma \begin{bmatrix} E & E & \cdots & E & E \\ E & E & & E & E \\ \vdots & & \ddots & & \vdots \\ E & E & & E & E \\ E & E & \cdots & E & E \end{bmatrix} \right\}$$

where $$D = \begin{bmatrix} 1 & -1 & p_2 \\ -1 & 1 & p_1 \\ p_2 & p_1 & -p_1 p_2 \end{bmatrix}$$ and $$\gamma = \frac{-(m_1 - m_2)^2}{(p_1+p_2) + (2N+1)(m_1 - m_2)^2}$$

$$\begin{bmatrix} a_1(\tau \neq 0) \\ a_2(\tau \neq 0) \\ -\lambda(\tau \neq 0) \end{bmatrix} = \frac{1}{p_1+p_2} \begin{bmatrix} \gamma_E \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{p_1+p_2} \begin{bmatrix} \beta\gamma \\ -\beta\gamma \\ \beta^2\gamma \end{bmatrix}$$

$$\begin{bmatrix} a_1(0) \\ a_2(0) \\ -\lambda(0) \end{bmatrix} = \frac{1}{p_1+p_2} \begin{bmatrix} D + \gamma E \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{p_1+p_2} \begin{bmatrix} p_2 + \beta\gamma \\ p_1 - \beta\gamma \\ -p_1p_2 + \beta^2\gamma \end{bmatrix}$$

The mean square error and noise power output are equal to $\lambda(0)$.

$$\lambda(0) = \frac{p_1p_2 - \beta^2\gamma}{p_1+p_2} = \frac{1}{p_1+p_2} \left\{ p_1p_2 + \frac{(p_2m_1 + p_1m_2)^2}{\left[(p_1+p_2)+(2N+1)(m_1-m_2)^2\right]} \right\}$$

Under the assumption that the white noise levels are the same for both channels, $p_1 = p_2 = p$ and the mean square error is

$$\lambda(0) = \frac{p}{2} \left[1 + \frac{(m_1 + m_2)^2}{2p + (2N+1)(m_1-m_2)^2} \right]$$

The noise output power for a beamsteer system applied to the same noise field is

$$\frac{1}{4}\left[(p_1+p_2) + (m_1+m_2)^2\right]$$

or

$$\frac{p}{2} \left[1 + \frac{(m_1 + m_2)^2}{2p} \right]$$

if $p_1 = p_2 = p$. Since the quantity $(2N+1)(m_1 - m_2)^2$ is always non-negative, the maximum-likelihood filter set reduces the noise output power to a level at least as low as the noise output power of a beamsteer system. Implicit in the design equations, however, is the assumption that "power" due to DC bias is as

undesirable as power due to fluctuations in the traces. In practice, the analyst easily "filters out" DC bias when visually inspecting a seismic trace.

Some important effects occur as a result of the means in the traces. The filter weights are

$$a_1(\tau \neq 0) = \frac{\beta\gamma}{p_1+p_2}$$

$$a_2(\tau \neq 0) = \frac{-\beta\gamma}{p_1+p_2}$$

$$a_1(0) = \frac{p_2}{p_1+p_2} + \frac{\beta\gamma}{p_1+p_2}$$

$$a_2(0) = \frac{p_1}{p_1+p_2} - \frac{\beta\gamma}{p_1+p_2}$$

If $p_1 = p_2 = p$ and α is set to $\beta\gamma/(p_1 + p_2)$, the frequency response of the filters is

$$A_1(f) = \frac{1}{2} + \alpha \sum_{\ell=-N}^{N} e^{-i2\pi f \ell \Delta t}$$

and

$$A_2(f) = \frac{1}{2} - \alpha \sum_{\ell=-N}^{N} e^{-i2\pi f \ell \Delta t}$$

If $\sigma(f)$ is set to $\sum_{\ell=-N}^{N} e^{-i2\pi f \ell \Delta t}$, the noise output power density of the maximum-likelihood filter set is

$$\left[A_1^*(f) A_2^*(f)\right] \begin{bmatrix} P(f) & 0 \\ 0 & P(f) \end{bmatrix} \begin{bmatrix} A_1(f) \\ A_2(f) \end{bmatrix}$$

$$= \left[\frac{1}{4} + \alpha\sigma(f) + \alpha^2\sigma^2(f)\right] P(f)$$

$$+ \left[\frac{1}{4} - \alpha\sigma(f) + \alpha^2\sigma^2(f)\right] P(f)$$

$$= \frac{1}{4}\left[P(f) + P(f)\right] + a\sigma(f)\left[P(f) - P(f)\right] + a^2\sigma^2(f)\left[P(f) + P(f)\right]$$

$$= \frac{P(f)}{2}\left[1 + 4a^2\sigma^2(f)\right]$$

as compared with $P(f)/2$ noise output power density for a beamsteer system. Thus, at any frequency other than DC, the beamsteer system reduces the noise level at least as much as the maximum-likelihood filter set. In fact

$$\sigma(f) = \sum_{\ell=-N}^{N} e^{-i2\pi f \ell \Delta t} = \frac{\sin(2N+1)\pi f \Delta t}{\sin \pi f \Delta t}$$

and $$\sigma(f) = 0 \quad \text{when } f = \frac{k}{\Delta t(2N+1)} \qquad (k = 1, 2, \ldots 2N)$$

The first zero of this function occurs at $f = 1/(2N+1)\Delta t$, and the function climbs monotonically until it approaches the value $2N+1$ as the frequency approaches zero from the right. The quantity $(2N+1)\Delta t$ is the effective length of the filter. This result provides a possible explanation for the poor performance of the maximum-likelihood filter set at frequencies below the reciprocal of the filter length.

C. TECHNIQUES USED TO ELIMINATE DC BIAS FROM THE DATA

The first step in removing DC bias from the traces was to apply single-channel prefilters to each of the site traces after rotation to the vertical trace. A 31-point convolution filter was used for this purpose. The weights for this filter are shown in Table II-1. Data points are scaled by 2^{-15} before exiting from the convolution filter microcode which implements the filter. The frequency response (after scaling) is shown in Figure II-1. The fact that the filter weights sum to zero guarantees that the response at DC is exactly zero or $-\infty$ in dB.

$$= \frac{1}{4}\left[P(f) + P(f)\right] + a\sigma(f)\left[P(f) - P(f)\right] + a^2\sigma^2(f)\left[P(f) + P(f)\right]$$

$$= \frac{P(f)}{2}\left[1 + 4a^2\sigma^2(f)\right]$$

as compared with $P(f)/2$ noise output power density for a beamsteer system. Thus, at any frequency other than DC, the beamsteer system reduces the noise level at least as much as the maximum-likelihood filter set. In fact

$$\sigma(f) = \sum_{\ell=-N}^{N} e^{-i2\pi f \ell \Delta t} = \frac{\sin(2N+1)\pi f \Delta t}{\sin \pi f \Delta t}$$

and $$\sigma(f) = 0 \text{ when } f = \frac{k}{\Delta t(2N+1)} \qquad (k = 1, 2, \ldots 2N)$$

The first zero of this function occurs at $f = 1/(2N+1)\Delta t$, and the function climbs monotonically until it approaches the value $2N+1$ as the frequency approaches zero from the right. The quantity $(2N+1)\Delta t$ is the effective length of the filter. This result provides a possible explanation for the poor performance of the maximum-likelihood filter set at frequencies below the reciprocal of the filter length.

C. TECHNIQUES USED TO ELIMINATE DC BIAS FROM THE DATA

The first step in removing DC bias from the traces was to apply single-channel prefilters to each of the site traces after rotation to the vertical trace. A 31-point convolution filter was used for this purpose. The weights for this filter are shown in Table II-1. Data points are scaled by 2^{-15} before exiting from the convolution filter microcode which implements the filter. The frequency response (after scaling) is shown in Figure II-1. The fact that the filter weights sum to zero guarantees that the response at DC is exactly zero or $-\infty$ in dB.

TABLE II-1

PRE-FILTER WEIGHTS

Lag	Weight
±15	-1384
±14	-1209
±13	-1049
±12	- 916
±11	- 819
±10	- 766
± 9	- 756
± 8	- 788
± 7	- 856
± 6	- 951
± 5	-1061
± 4	-1174
± 3	-1277
± 2	-1360
± 1	-1413
0	31558

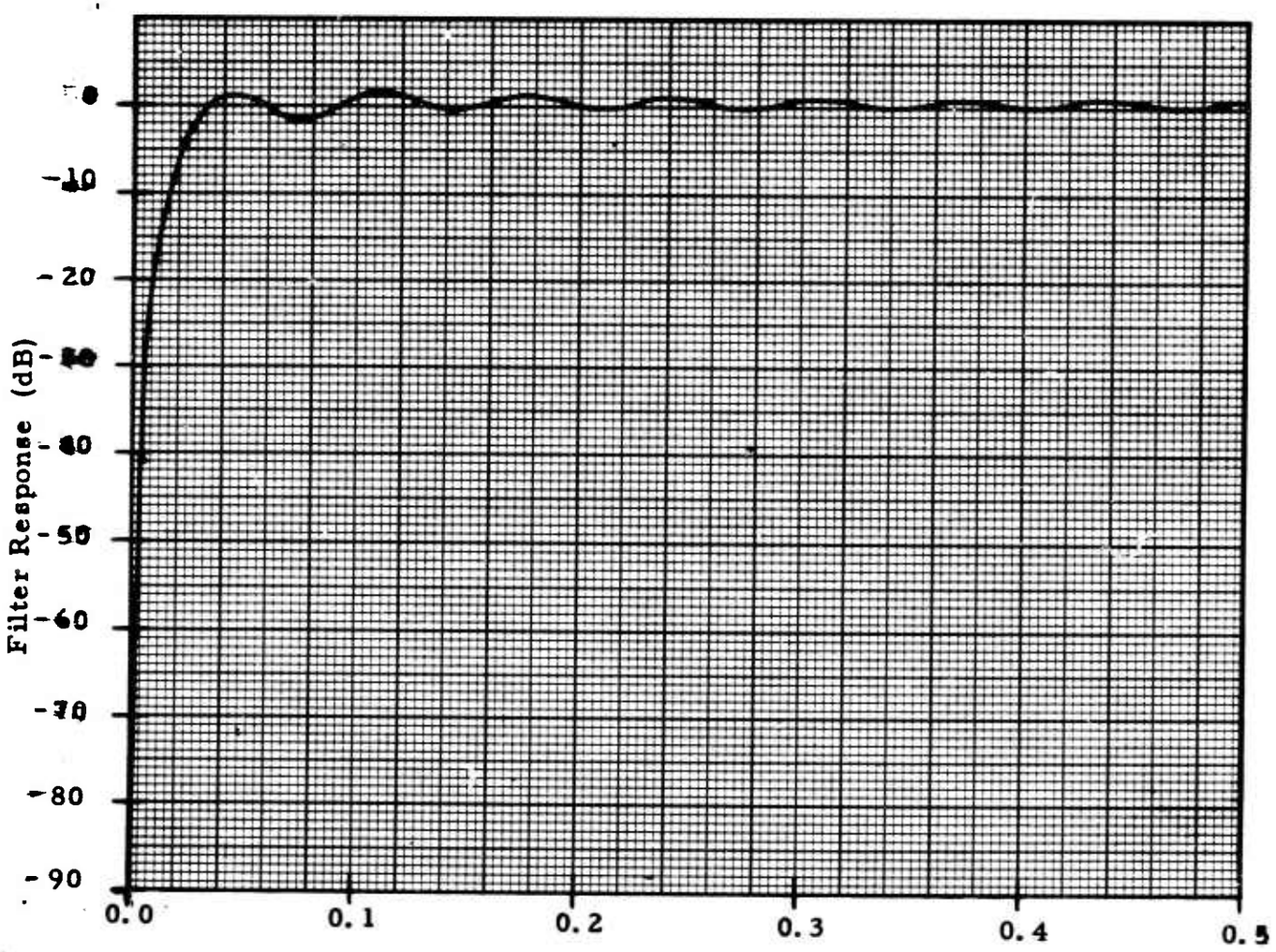


FIGURE II-1

PRE-FILTER RESPONSE (dB)

The filter of Table II-1 and Figure II-1 was obtained from a computer program which minimizes, for a discrete convolution filter of specified length, the weighted mean square error (integrated over the Nyquist band) between the desired frequency response and the convolution-filter frequency response. After the initial run, several iterations with altered frequency weightings and desired responses were made to reduce the ripple in the filter response. Next the desired response at 0 Hz was modified so as to reduce the amplitude at DC. After a response below -100 dB was achieved at 0 Hz, the filter coefficients were rounded to the nearest 1/32,768. Finally, the zero lag weight of the convolution filter was adjusted by a multiple of 1/32,768 so that the sum of the filter coefficients was zero (yielding a response of $-\infty$ dB at DC).

When the program changes to implement this filter were made, difficulties persisted. A detailed study of the adaptive-filtering subroutine showed that 1/2 count of negative bias per channel was being introduced by the convolution-filter microcode because of truncation. The subroutine was rewritten to compensate for this bias in the beam output trace and to round intermediate results wherever possible. After these modifications, the DC bias problems were eliminated.

SECTION III
EFFECT OF ROUNDOFF ERROR IN THE FILTER UPDATE EQUATION

A. DISCUSSION

In the adaptive-filtering investigations conducted for this study, the bulk of the adaptive-filter update calculations and the adaptive filter output computations were performed using integer arithmetic. The reason for integer arithmetic was the desirability of operating in an on-line mode on the IBM 360/40 computers at SDAC using their special high-speed microcode operations. In the construction of hardware specifically tailored to implement adaptive beamforming, economic considerations seem to dictate that integer arithmetic be used. For this reason, the experience gained during this study in the problems arising from integer-arithmetic roundoff error may well be applicable to a variety of economically practical adaptive-filtering systems.

The objective of this section is to obtain quantitative estimates of the directional error in the filter update vector as a function of the limited number of parameters controlling the directional error. Some of these parameters, namely the number of bits used to represent the input data points and the filter-weight coefficients, are subject to direct human control. Other parameters depend either directly or indirectly on the characteristics of the data at any given array. Once directional-error tolerances can be specified with sufficient precision to avoid significant degradation of adaptive-filter performance, noise field measurements at an array, together with the directional-error approximations given in this section, should provide useful estimates of the number of bits needed for the input data points and the filter-

weight coefficients. Even estimates accurate only to the nearest four bits are valuable: it is extremely useful, in designing adaptive-filtering hardware, to know whether 12 bits, 16 bits, 20 bits, 24 bits, etc., are needed to represent adequately the data and adaptive-filter coefficients.

In this section, the computational error will be traced through the series of integer-arithmetic operations performed in updating the adaptive-filter weights. At this point, these operations will be reviewed. The filter update procedure is a vector operation which can be written compactly in the vector equation

$$A^{new} = A^{old} + \frac{2K_s X^T A^{old} (\bar{X}-X)}{(\bar{X}-X)^T (\bar{X}-X)} ,$$

which describes the particular algorithm used in this study. The term K_s is a scalar quantity called the convergence factor. It controls the adaptation rate of the adaptive-filtering process. The dot product $X^T A^{old}$ denotes the series of multiply-and-add operations which yield the adaptive beamformer output $y(t)$. The superscript T denotes vector transposition. The vectors A^{new}, A^{old}, $\bar{X}$, and X (as well as their constituent components) were defined previously in subsections I-D and I-E. In the algorithm actually employed, the denominator $(\bar{X}-X)^T(\bar{X}-X)$ is approximated by

$$(2N+1) \sum_{i=1}^{M} P_i(t) ,$$

where $P_i(t)$ is a moving power average of the difference between the i-th channel and the beamsteer output (see page I-8). In this approximation, M is the number of input channels and $(2N+1)$ is the total number of filter weights per channel.

Since the denominator $(\bar{X}-X)^T(\bar{X}-X)$ and the convergence factor K_s are positive scalar quantities, the direction of motion of the filter update vector $(A^{new} - A^{old})$ is determined by the sign of the adaptive filter output $y(t)$, a scalar quantity, and the vector $(\bar{X}-X)$. Directional error in the vector $(\bar{X}-X)$ contributes to directional error in the filter update vector. Errors in the data vector X, moreover, may change the sign of the adaptive filter output $y(t) = X^T A^{old}$ or cause it to round to zero: in the first case, the filter vector moves in a direction which tends to increase the average squared filter output $\overline{y^2(t)}$; in the second case, the adaptive filter vector cannot move. After multiplication of the vector $(\bar{X}-X)$ by the scalar quantity

$$\frac{2K_s X^T A^{old}}{(\bar{X}-X)^T(\bar{X}-X)} ,$$

the individual components of the resultant scaled vector are rounded to the nearest integer in the numerical representation of the filter weights and then added to the corresponding components of the old filter vector. The filter-weight roundoff process just described introduces further directional error into the filter update vector. The filter-weight roundoff error can be made as small as desired by using a sufficient number of bits in the numerical representation of the filter weights. A similar capability, however, is not possible in computing the adaptive filter output $y(t)$, where the preservation of a non-zero value with the correct sign is ultimately limited by the quantization error introduced by the digitization system.

The principal questions to be answered in determining the error in the direction of movement of the filter weight vector are, therefore,

- How much directional accuracy is maintained in the vector $(\bar{X}-X)$ after machine computational procedures have been performed?

- How often do computational and digitization errors reverse the sign of the adaptive filter output or result in a zero value for the adaptive filter output when it is rounded to the nearest integer?

- What precision is required in the numerical representation of the filter weights to achieve acceptable error levels when the filter update vector is added to the old filter weight vector?

Subsections B and C examine error in the vector $(\overline{X}-X)$. Subsection B, a necessary prelude to subsection C, discusses errors in a single component of the vector $(\overline{X}-X)$. Error is traced through the individual processes leading to the final value $\overline{x}(t-j) - x_i(t-j)$. Probability densities, standard deviations of error, and maximum errors are presented for rotated data, prefiltered data, the beamsteer output, and the beamsteer output minus a single channel.

After the preliminaries of subsection B, directional error in the vector $(\overline{X}-X)$ is investigated in subsection C. The error E in $(\overline{X}-X)$ is separated into a part E_c satisfying the filter-weight constraint conditions and a part E_b perpendicular to the vector space corresponding to the filter-weight constraints. Probability distributions for the angle of error in $(\overline{X}-X)$ and the angle of error in its projection onto the constraint space are calculated.

Subsection D considers the effect of digitization and roundoff error on the adaptive filter output $y(t)$. The probability that $y(t)$ rounds to zero and the probability that $y(t)$ changes sign are estimated. Since these probabilities depend on the filter vector A, a derivation of the maximum-likelihood filter-design equations is included so that the filter vector may be determined from the data crosscorrelation matrix.

Subsection E studies the consequences of rounding off the updated filter vector A^{new}. To evaluate the likelihood that the roundoff process

immobilizes the filter vector by preventing any change in the individual components, an expression for the squared magnitude $|A^{new} - A^{old}|^2$ is employed. To estimate the mean angle of error in the vector $A^{new} - A^{old}$, the probability distributions associated with rounding off the filter weights are determined. Finally, two different techniques for eliminating discrepancies in the maximum-likelihood constraints are appraised in terms of the resultant directional error in the filter update vector.

Subsection F is a summary of the results in this section.

B. ERRORS ASSOCIATED WITH A SINGLE COMPONENT OF THE VECTOR $(\bar{X}-X)$

As a first step in estimating the directional error of the vector $(\bar{X}-X)$, the errors associated with a single component must be established. In order to evaluate the directional error, it is necessary to distinguish between error in one component of $(X-X)$ due to inaccuracies in the quantities $x_i(t-j)$ prior to forming the beamsteer output and error arising solely from roundoff error during computation of the beamsteer output $\bar{x}(t-j)$ for time $t - j\Delta t$. This distinction is necessary because error prior to generating the beamsteer output cannot dislodge the updated filter vector A^{new} from the subspace corresponding to the maximum-likelihood constraints. Conversely, roundoff error in calculating the beamsteer output forces motion perperdicular to the constraint space. A detailed description of this phenomenon will be given at a more appropriate point in this section.

In analyzing errors at the single-component level, the digitization process is presumed to introduce an error with a uniform probability density between plus and minus one-half count. There is one realistic situation where this premise is not justified. The ALPA data values are represented as 16-bit gain-ranged numbers with a 12-bit, two's-complement fraction and a four-bit negative exponent. When the DC level on any triax component rises to 2048

computer counts or more, the digitization error is effectively multiplied by the appropriate power of two. For the data samples processed in this report, one or two of the six channels input to the adaptive processor might have floating means of this size. To achieve simplicity of presentation, this fact is ignored. In neglecting this possibility, the digitization error is underestimated. This underestimation should not materially affect the results obtained in this section.

A different assumption is made for computational roundoff error. When an intermediate result must be rounded to the nearest count, the resultant roundoff error is assumed to lie between -1/2 and +1/2 at equally-spaced discrete points centered about zero. If none of these discrete points is located one-half count from zero, they are assumed to equally likely. If an error of one-half count is possible, each of the end points is assumed to be half as likely as each of the interior points. Up to and including the point where the adaptive channel data points $x_i(t-j)$ are averaged to form the beamsteer output, the error of each quantity involved in a summation is assumed to be independent of the error in the other quantities involved in the summation. In the case of independent summand errors, the probability density function for the error in the sum is the convolution of the probability density function for the error due to roundoff with that of each individual summand error.

When the possibility that roundoff error could be affecting adaptive filter performance was first considered, an experiment was conducted. First, with the convergence factor K_s set to 0.30, a noise sample from a relatively quiet summer day was run through the adaptive filter program. Noise in the adaptive filter output was reduced by 2 dB relative to the beamsteer output. Second, the data points were multiplied by 16 before processing. At the same convergence rate, noise reduction was approximately 6 dB. It was apparent, therefore, that roundoff error was significantly affecting processor

performance. Scaling the data by a factor of 16 was sufficient to achieve noise reduction comparable to that obtained for the same noise sample by a separate computer program using floating-point arithmetic.

In this section, error will be examined for the case of unscaled data, data scaled by a factor of 16, and data where the floating-point arithmetic is used in the adaptive processor. When floating-point arithmetic is employed, roundoff error is assumed to be negligible, and error control is ultimately limited by digitization accuracy. A comparison of the error in unscaled data and data scaled by a factor of 16 shows the error reduction achieved by scaling the data. A comparison of the error in the scaled data and the data processed using floating-point arithmetic indicates how closely the scaled data approaches ultimate achievable precision.

In the first stage of processing, three-component triax data at each site are combined to form a vertical component by multiplying each triax component by one third and summing all three components (rounding to the nearest count when the vertical component is output). In the case of unscaled data, the error in each summand has a uniform probability density between plus and minus one sixth (from digitization error), and a roundoff error with equal probability at the three discrete points $-1/2$, 0, and $1/3$. Since the probability density function is the convolution of four separate probability density functions, the resultant probability density function has a Fourier transform which is the product of the individual transforms for each of the four separate independent probability density functions. For a uniform density between $-c/2$ and $c/2$, the Fourier transform is

$$\int_{-c/2}^{c/2} \frac{e^{-i2\pi sy}}{c}\, dy = \frac{e^{-i\pi cs}}{-i2\pi cs} - \frac{e^{i\pi cs}}{-i2\pi cs}$$

$$= \frac{\sin \pi cs}{\pi cs} \quad (\text{or } 1 \text{ if } s = 0)$$

$$= \operatorname{sinc} cs,$$

where

$$\operatorname{sinc} s = \frac{\sin \pi s}{\pi s} \quad (\text{or } 1 \text{ if } s = 0)$$

For a finite-comb probability density function

$$p(y) = \frac{1}{n} \sum_{j=-\frac{(n-1)}{2}}^{\frac{n-1}{2}} \delta\left(y - \frac{j}{n}\right),$$

where n is an odd integer and δ is the Dirac delta function, the Fourier transform is

$$P(s) = \int_{-\infty}^{\infty} p(y)\, e^{-i2\pi sy}\, dy = \frac{1}{n} \sum_{j=-\frac{(n-1)}{2}}^{\frac{n-1}{2}} e^{\frac{-i2\pi js}{n}}.$$

The transform is a geometric progression, so that

$$\left(e^{\frac{i\pi s}{n}} - e^{\frac{-i\pi s}{n}}\right) P(s) =$$

$$\frac{1}{n}\left[\sum_{j=-\frac{n}{2}}^{\frac{n}{2}-1} e^{-\frac{i2\pi js}{n}} - \sum_{j=-(\frac{n}{2}-1)}^{\frac{n}{2}} e^{-\frac{i2\pi js}{n}}\right]$$

$$= \frac{1}{n}\left[e^{i\pi s} - e^{-i\pi s}\right]$$

and

$$P(s) = \frac{1}{n}\left[\frac{\left(e^{i\pi s} - e^{-i\pi s}\right)/2i}{\left(e^{\frac{i\pi s}{n}} - e^{\frac{-i\pi s}{n}}\right)/2i}\right]$$

$$= \frac{\sin\pi s}{n \sin(\pi s/n)} \qquad \text{(or 1 if } s \text{ is an integer multiple of } n\text{).}$$

If the function $\mathcal{H}_{11}(s)$ is defined for odd integers as the Fourier transform just computed for the finite-comb probability density function, the Fourier transform of the error in the vertical component for unscaled data is $\text{sinc}^3(s/3)\,\mathcal{H}_3(s)$. The maximum possible error is 5/6 count. In the case of data scaled by a factor of 16, the digitization error is multiplied by 16 during scaling, but roundoff error is still the same. The Fourier transform of the ensuing error probability density function is $\text{sinc}^3(16s/3)\,\mathcal{H}_3(s)$. The maximum possible error is 8 1/3 counts. (At the same time that the absolute error is higher because of scaling, the data points are 16 times larger, so that the relative error is reduced.) For data going through a floating-point processor, there is negligible roundoff error. The Fourier transform of the corresponding probability density function is $\text{sinc}^3(s/3)$. The maximum possible error is 1/2 count. Inverse Fourier transforms graphing the error probability density per count for unscaled data, data scaled by 16, and data

passing through a floating-point processor are presented in Figure III-1. The physical size of the vertical scale was increased by 16 and that of the horizontal scale reduced by 16 on the page in the case of data scaled by 16 to permit meaningful comparison between the three modes of processing for the relative error in the vertical-component output.

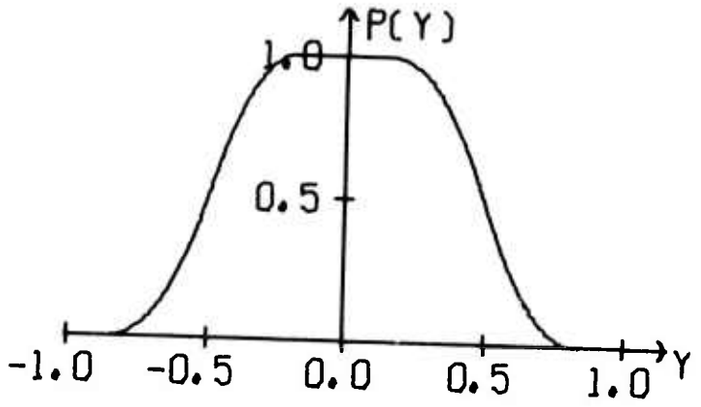
In the second stage of processing, the vertical-component data traces are funneled through a filter with exactly zero response at DC. The filter weights (scaled by 2^{15}) for this filter were given previously on page II-9. The equation which describes the filter implementation is

$$x_i(t) = \sum_{j=-15}^{15} b_j \, v_i(t - j) ,$$

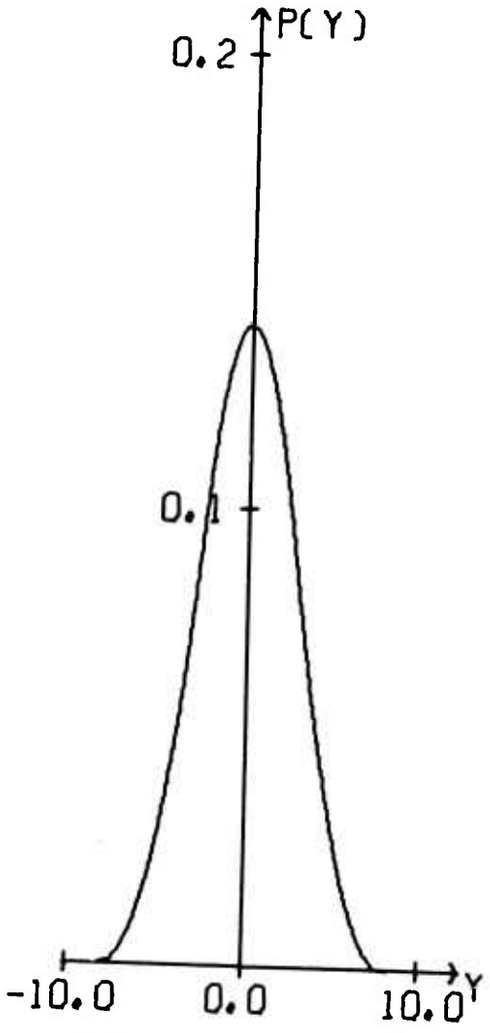
where $x_i(t)$ is the prefiltered vertical-component trace for the i-th site at time t, b_j is the filter weight for the j-th lag, and $v_i(t - j)$ is the vertical-component input trace for the i-th site at time $t - j\Delta t$. Since the filter weights are all multiples of 2^{-15}, the roundoff error is presumed to lie at any of the 32769 equally-spaced points between plus and minus one-half count. Both of the end points are presumed to be half as likely as any of the interior points, so that the probability density function for the roundoff error is

$$p(y) = \frac{1}{2n} \delta(y + 1/2) + \frac{1}{n} \left[\sum_{j=-(\frac{n}{2}-1)}^{\frac{n}{2}-1} \delta(y - \frac{j}{n}) \right] + \frac{1}{2n} \delta(y-1/2)$$

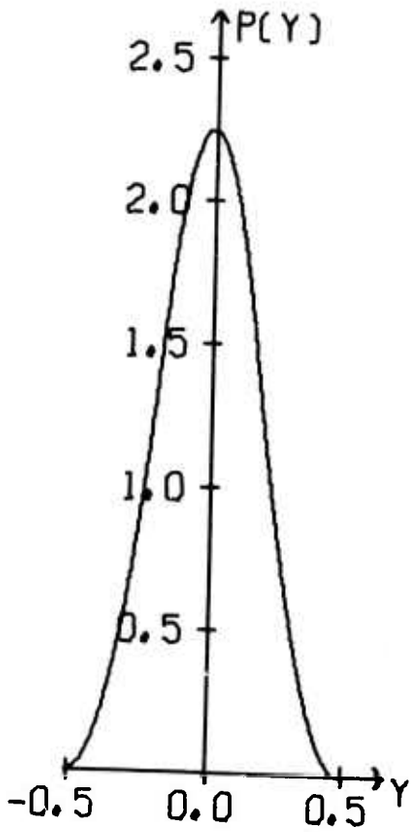
$$= \frac{1}{2n} \left[\sum_{j=-\frac{n}{2}}^{\frac{n}{2}} \delta(y - \frac{j}{n}) + \sum_{j=-(\frac{n}{2}-1)}^{\frac{n}{2}-1} \delta(y - \frac{j}{n}) \right] ,$$



(a) Unscaled Data



(b) Data Scaled by 16



(c) Floating-Point Data

FIGURE III-1

ERROR PROBABILITY DENSITY PER COUNT AFTER FORMATION OF VERTICAL COMPONENT

where n is 32768. The corresponding Fourier transform is

$$P(s) = \int_{-\infty}^{\infty} p(y)\, e^{-i2\pi sy}\, dy$$

$$= \frac{1}{2n}\left[\sum_{j=-\frac{n}{2}}^{\frac{n}{2}} e^{-\frac{i2\pi js}{n}} + \sum_{j=-(\frac{n}{2}-1)}^{\frac{n}{2}-1} e^{-\frac{i2\pi js}{n}}\right],$$

which is the sum of two geometric progressions. Hence

$$\left(e^{\frac{i\pi s}{n}} - e^{-\frac{i\pi s}{n}}\right) P(s)$$

$$= \frac{1}{2n}\left[e^{\frac{i\pi(n+1)s}{n}} + e^{\frac{i\pi(n-1)s}{n}} - e^{-\frac{i\pi(n-1)s}{n}} - e^{-\frac{i\pi(n+1)s}{n}}\right]$$

$$= \frac{1}{2n}\left(e^{\frac{i\pi s}{n}} + e^{-\frac{i\pi s}{n}}\right)\left(e^{i\pi s} - e^{-i\pi s}\right),$$

so that

$$P(s) = \frac{1}{n}\left(\frac{e^{\frac{i\pi s}{n}} + e^{-\frac{i\pi s}{n}}}{2}\right)\left[\frac{\left(e^{i\pi s} - e^{-i\pi s}\right)/2i}{\left(e^{\frac{i\pi s}{n}} - e^{-\frac{i\pi s}{n}}\right)/2i}\right]$$

$$= \left[\frac{\sin\pi s}{n \sin(\pi s/n)}\right] \cos\left(\frac{\pi s}{n}\right),$$

or 1 when s is an integer multiple of n. The function $\mathcal{H}_n(s)$ for even integers is defined as the expression just given. Under the assumption that the errors in the separate vertical-component input points are mutually independent and, further, independent of the roundoff error at output, the Fourier transform of the resulting error probability density function for unscaled data is

$$\mathcal{H}_{32768}(s) \prod_{j=-15}^{15} \operatorname{sinc}^3\left(\frac{b_j s}{3}\right) \mathcal{H}_3(b_j s).$$

Since the prefilter is symmetric about zero, $b_{-j} = b_j$ and the Fourier transform above may be written

$$\mathcal{H}_{32768}(s) \operatorname{sinc}^3\left(\frac{b_o s}{3}\right) \mathcal{H}_3(b_o s) \prod_{j=1}^{15} \operatorname{sinc}^6\left(\frac{b_j s}{3}\right) \left[\mathcal{H}_3(b_j s)\right]^2 .$$

The maximum possible error for unscaled data is

$$\frac{1}{2} + \frac{5}{6}\left(\sum_{j=-15}^{15} |b_j|\right)$$

counts or slightly less than 2 1/6 counts since the sum of the filter weight absolute values is slightly less than 2 . For data scaled by 16, the Fourier transform of the error probability density function is

$$\mathcal{H}_{32768}(s) \operatorname{sinc}^3\left(\frac{16 b_o s}{3}\right) \mathcal{H}_3(b_o s) \prod_{j=1}^{15} \operatorname{sinc}^6\left(\frac{16 b_j s}{3}\right)\left[\mathcal{H}_3(b_j s)\right]^2$$

and the maximum possible error is

$$\frac{1}{2} + \frac{25}{3}\left(\sum_{j=-15}^{15} |b_j|\right)$$

counts or slightly less than 17 1/6 counts. For data processed with floating-point arithmetic, the corresponding Fourier transform is

$$\operatorname{sinc}^3\left(\frac{b_o s}{3}\right) \prod_{j=1}^{15} \operatorname{sinc}^6\left(\frac{b_j s}{3}\right)$$

and the maximum possible error is

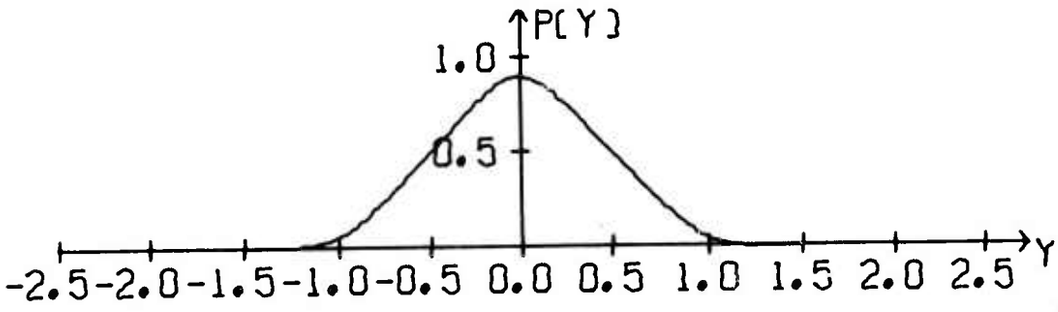
$$\frac{1}{2}\left(\sum_{j=-15}^{15}\left|b_j\right|\right)$$

counts or slightly less than 1 count. Inverse Fourier transforms giving the error probability density per count after prefiltering are shown in Figure III-2 for unscaled data, data scaled by 16, and data processed with floating-point arithmetic.

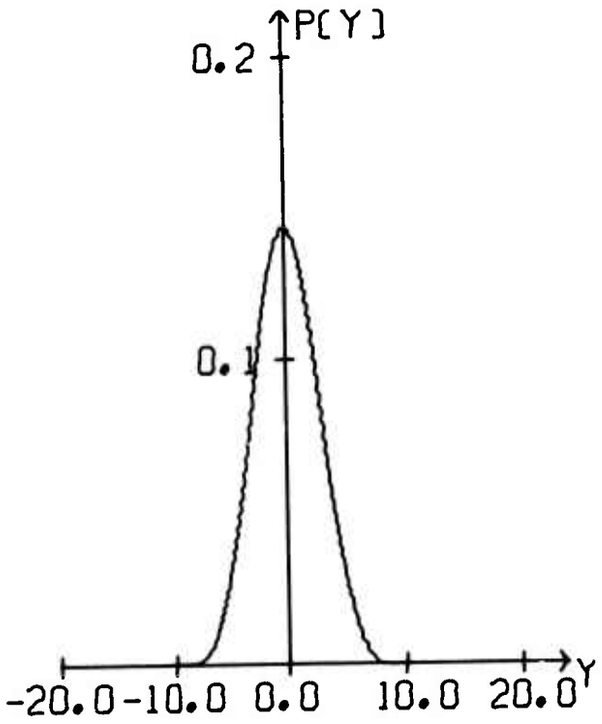
In the third stage of processing, the prefiltered vertical-component data traces for all sites at time t are summed and divided by the number of channels to create the beamsteer output. When division is performed, the result is rounded to the nearest count. Multiplication by the reciprocal of the number of channels has the effect of dividing the argument of the Fourier transforms in the previous stage by the number of channels. On the other hand, the assumption of mutually independent error between the individual prefiltered vertical components at each site means that the Fourier transforms corresponding to the sum of the traces (scaled by 1/M) must be raised to a power equal to the number of channels. Prior to rounding off to the nearest integer, therefore, the Fourier transforms of the beamsteer output error probability density functions are

$$\left[\mathcal{K}_{32768}\left(\frac{s}{6}\right)\right]^6 \operatorname{sinc}^{18}\left(\frac{b_o s}{18}\right)\left[\mathcal{K}_3\left(\frac{b_o s}{6}\right)\right]^6 \prod_{j=1}^{15} \operatorname{sinc}^{36}\left(\frac{b_j s}{18}\right)\left[\mathcal{K}_3\left(\frac{b_j s}{6}\right)\right]^{12},$$

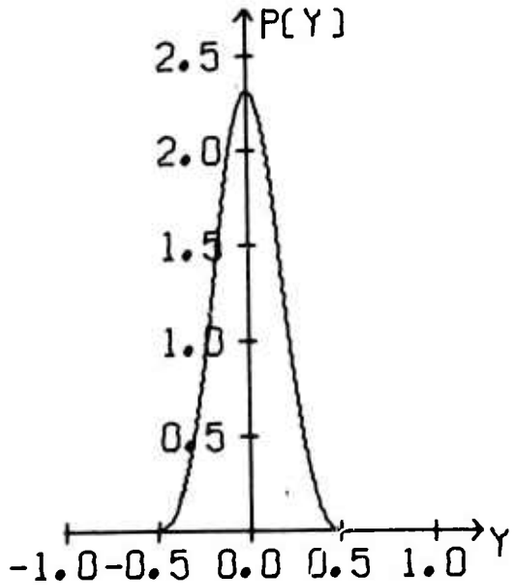
$$\left[\mathcal{K}_{32768}\left(\frac{s}{6}\right)\right]^6 \operatorname{sinc}^{18}\left(\frac{16 b_o s}{18}\right)\left[\mathcal{K}_3\left(\frac{b_o s}{6}\right)\right]^6 \prod_{j=1}^{15} \operatorname{sinc}^{36}\left(\frac{16 b_j s}{18}\right)\left[\mathcal{K}_3\left(\frac{b_j s}{6}\right)\right]^{12},$$



(a) Unscaled Data



(b) Data Scaled by 16



(c) Floating-Point Data

FIGURE III-2

ERROR PROBABILITY DENSITY PER COUNT AFTER PREFILTERING VERTICAL COMPONENT

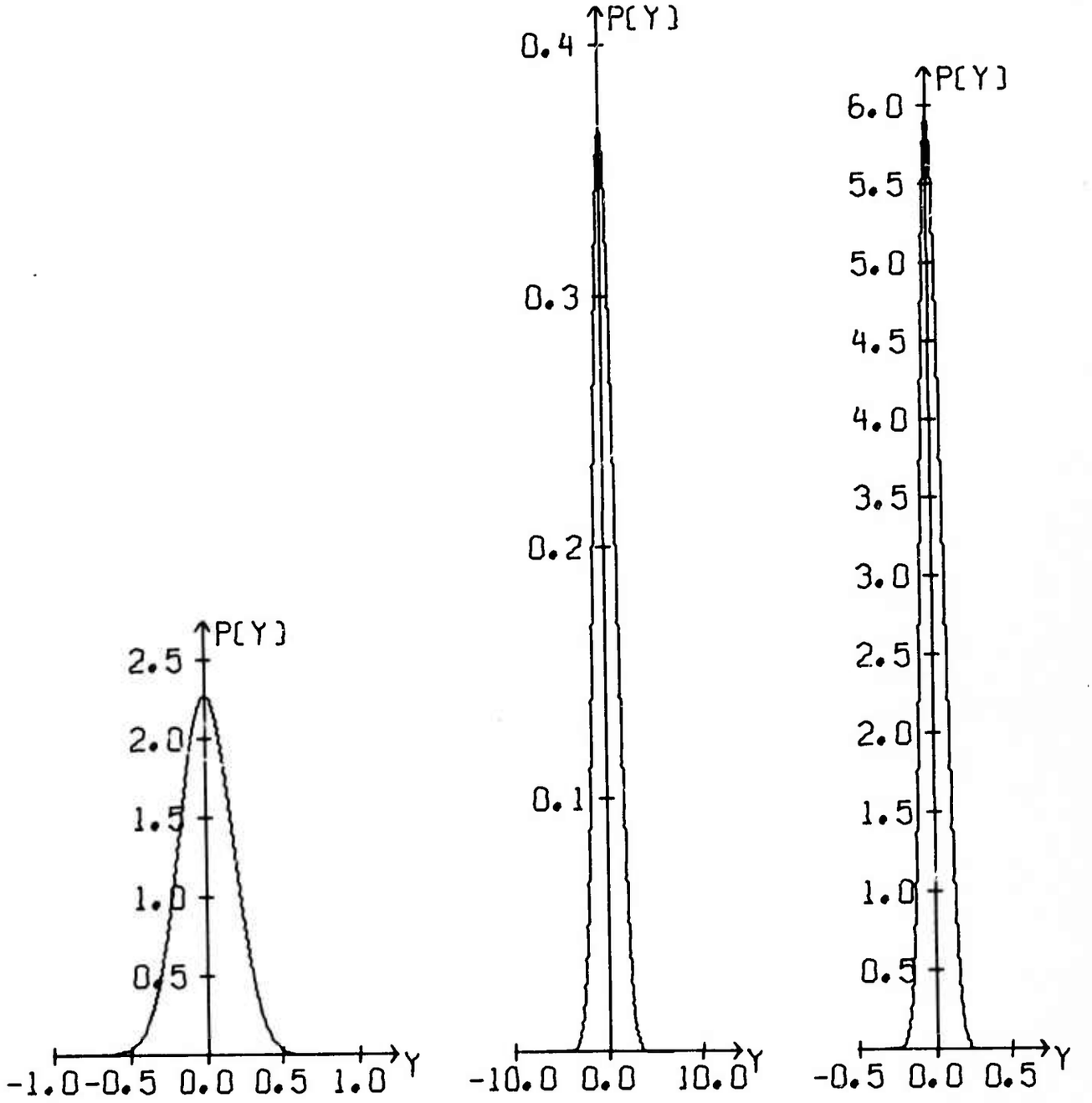
and

$$\text{sinc}^{18}\left(\frac{b_o s}{18}\right) \prod_{j=1}^{15} \text{sinc}^{36}\left(\frac{b_j s}{18}\right)$$

for unscaled data, data scaled by 16, and floating-point data, respectively, when six channels are used in the adaptive processor. The beamsteer error probability density functions prior to roundoff are shown in Figure III-3 for the three cases considered. Before rounding to the nearest integer, the maximum possible errors are exactly the same as in the previous stage -- 2 1/6 counts, 17 1/6 counts, and 1 count, respectively.

In those instances where integer arithmetic is employed, roundoff error may occur at the seven discrete points ranging from minus one-half to plus one-half count at increments of 1/6 count. The two end points are assumed to be half as likely as the interior points, so that the Fourier transform for the roundoff error is $\mathcal{K}_6(s)$ both for unscaled data and data scaled by a factor of 16. The beamsteer roundoff error probability density for both cases is pictured in Figure III-4. According to the assumptions stated earlier, floating-point arithmetic produces negligible roundoff error in the beamsteer output.

To determine the beamsteer error probability density functions after roundoff for unscaled data and data scaled by 16, the Fourier transforms prior to roundoff are multiplied by $\mathcal{K}_6(s)$, and inverse Fourier transforms are taken. The probability density function for the floating-point beamsteer error is the same as before. Figure III-5 gives the three specified probability density functions. After roundoff, the maximum possible errors are 2 2/3 counts, 17 2/3 counts, and 1 count, respectively. One-half count is added to the maximum beamsteer output error when the output is rounded to the nearest integer. No roundoff procedure is performed when floating-point arithmetic is used.



(a) Unscaled Data (b) Data Scaled by 16 (c) Floating-Point Data

FIGURE III-3

BEAMSTEER ERROR PROBABILITY DENSITY PER COUNT BEFORE ROUNDOFF

$$\frac{\delta(y+1/2)}{12} + \frac{\delta(y+1/3)}{8} + \frac{\delta(y+1/6)}{6} + \frac{\delta(y)}{6} + \frac{\delta(y-1/6)}{6} + \frac{\delta(y-1/3)}{6} + \frac{\delta(y-1/2)}{12}$$

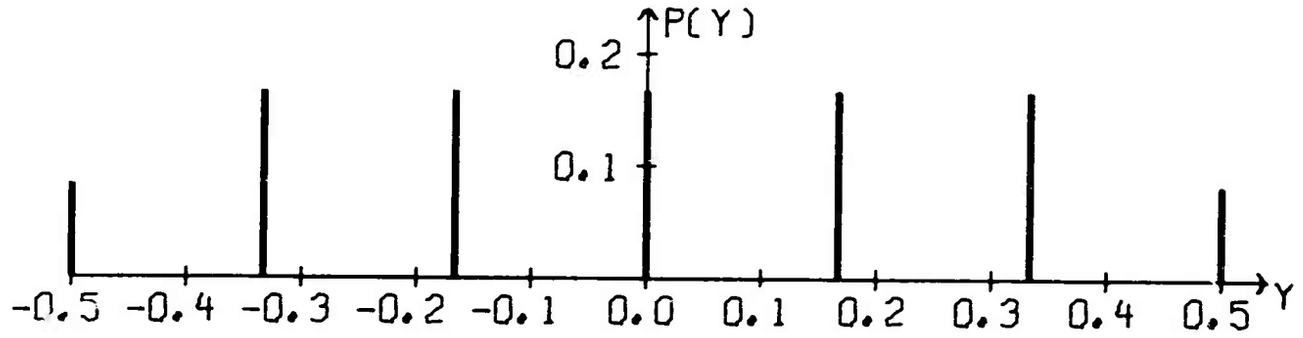
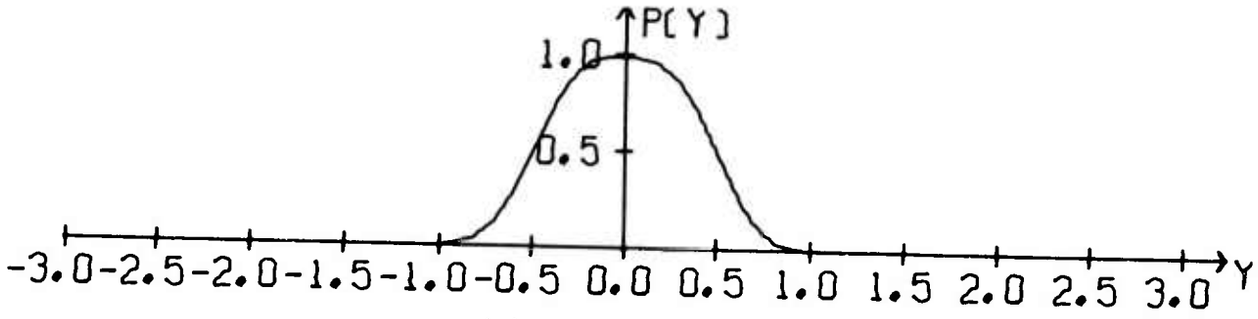
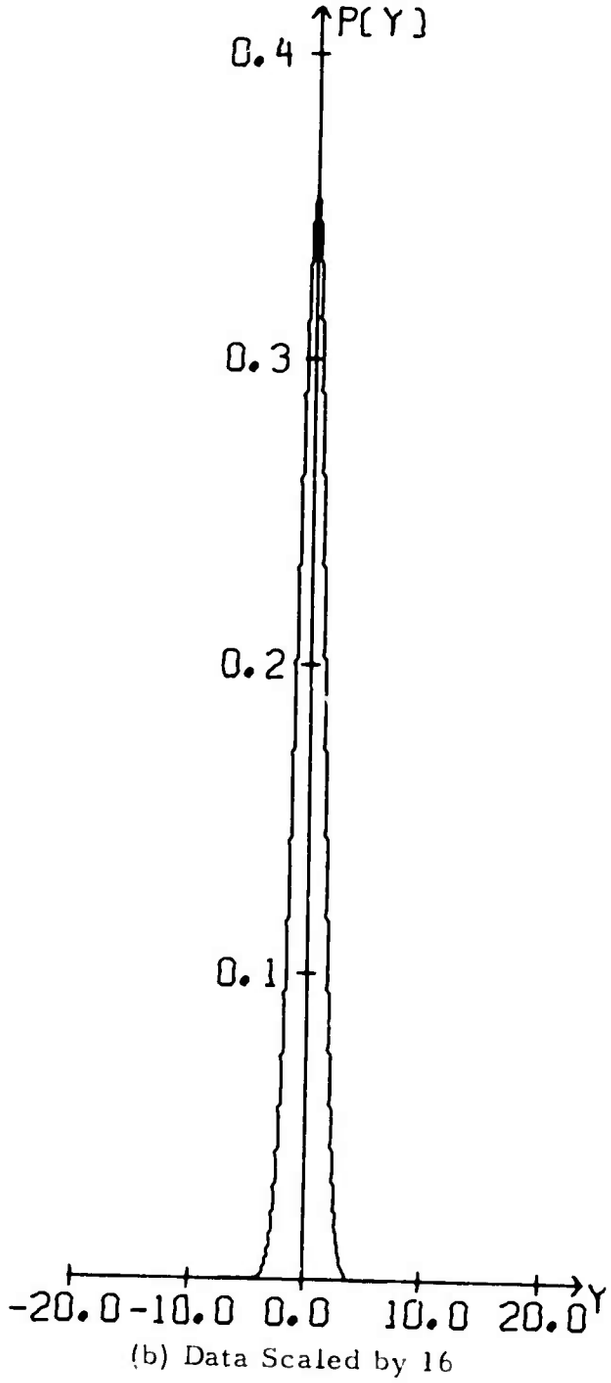


FIGURE III-4

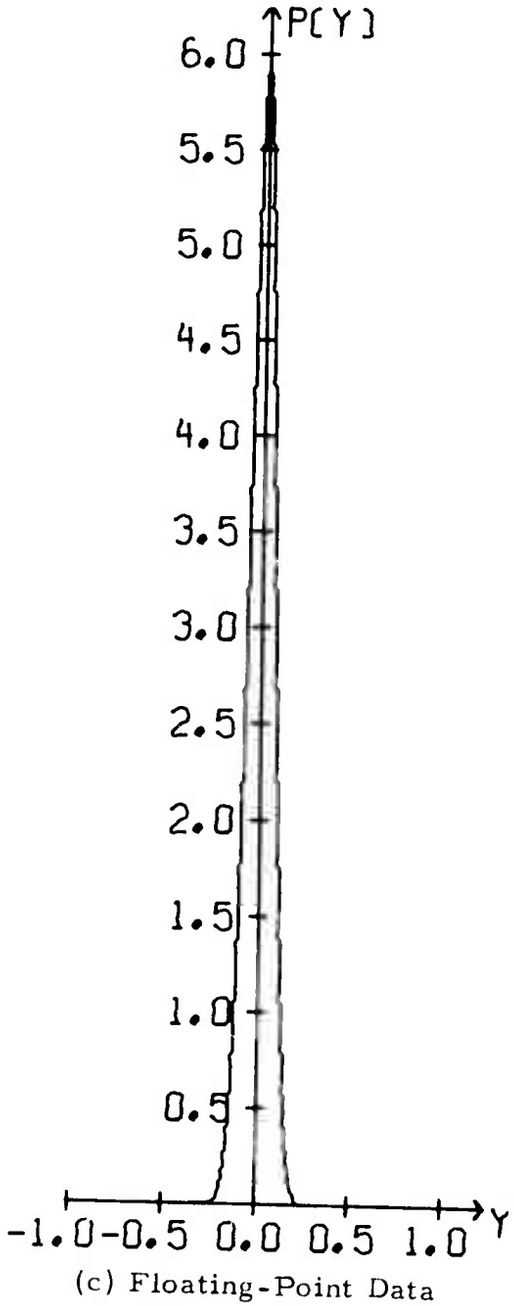
PROBABILITY DENSITY PER COUNT FOR
BEAMSTEER ROUNDOFF ERROR



(a) Unscaled Data



(b) Data Scaled by 16



(c) Floating-Point Data

FIGURE III-5

BEAMSTEER ERROR PROBABILITY DENSITY PER COUNT AFTER ROUNDOFF

In the fourth stage of processing, each individual component of the vector $(\bar{X} - X)$ is formed by subtracting the prefiltered vertical component for the i-th site from the beamsteer output at the same time. No further roundoff error is produced: either integers are subtracted from integers or floating-point arithmetic generates no significant error. The error in an individual component of $(\bar{X} - X)$ is

$$\frac{1}{M} \sum_{k=1}^{M} \left[x_k(t-j) + \epsilon_k(t-j) \right] + \epsilon_b(t-j) - \left[x_i(t-j) + \epsilon_i(t-j) \right]$$

$$- \left\{ \left[\frac{1}{M} \sum_{k=1}^{M} x_k(t-j) \right] - x_i(t-j) \right\}$$

$$= \epsilon_b(t-j) + \frac{1}{M} \left[\sum_{k=1}^{M} \epsilon_k(t-j) \right] - \epsilon_i(t-j)$$

$$= \epsilon_b(t-j) + \frac{1}{M} \left[\sum_{\substack{k=1 \\ k \neq i}}^{M} \epsilon_k(t-j) \right] - \frac{(M-1)}{M} \epsilon_i(t-j) ,$$

where $\epsilon_b(t-j)$ is the beamsteer output roundoff error at time $t-j\Delta t$ (due solely to the roundoff operation carried out at the tail end of the beamsteer output computation), where $\epsilon_i(t-j)$ is the error in the prefiltered vertical component for the i-th site at time $t-j\Delta t$ (the same component for which the error in $\bar{X} - X$ is being evaluated), where $\epsilon_k(t-j)$ is the error in the prefiltered vertical component of the k-th site at time $t-j\Delta t$, and where M is the number of channels. If the term $\epsilon_b(t-j)$ is ignored and P(s) is the Fourier transform of the probability density function for the error in any prefiltered vertical component, then

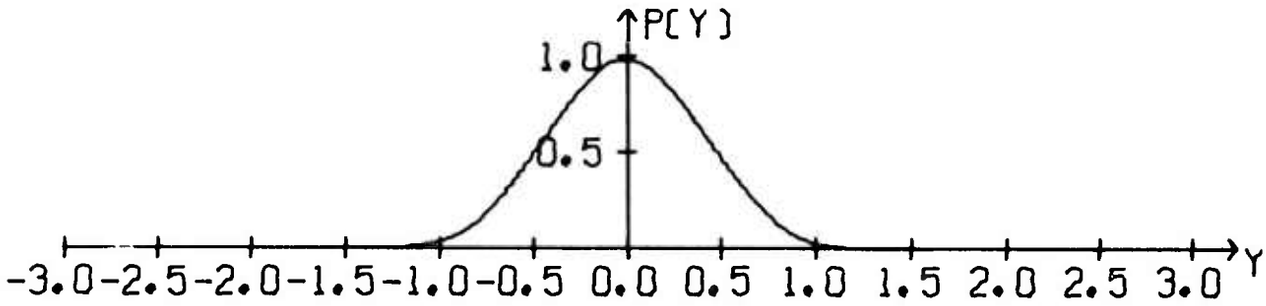
$$\left[P\left(\frac{s}{6}\right) \right]^5 \qquad P\left(\frac{5s}{6}\right)$$

is the Fourier transform of the probability density function for the error in each component of the vector $(\bar{X} - X)$. (As in the case where the beamsteer output error was discussed, the random variables $\epsilon_k(t-j)$ are considered to be mutually independent.) Figure III-6 shows the corresponding probability density functions for unscaled data, data scaled by 16, and floating-point data, respectively. In each case, the maximum possible error (ignoring the beamsteer output roundoff error) is 5/3 as large as the maximum possible prefiltered vertical-component error. As a result, the corresponding maximum errors are 3 11/18, 28 11/18, and 1 2/3 counts. When the term $\epsilon_b(t-j)$ is not ignored, the Fourier transform of the probability density function for the error in each component of the vector $(\bar{X} - X)$ is

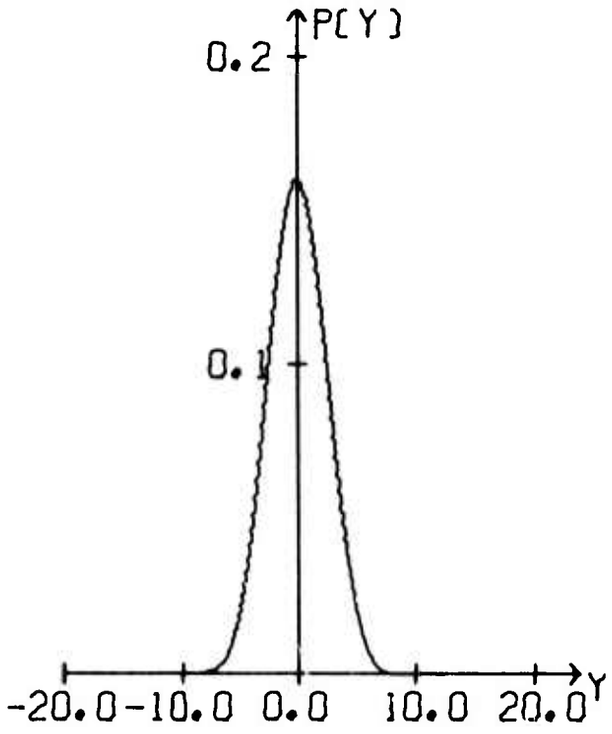
$$\mathcal{H}_6(s) \left[P\left(\frac{s}{6}\right)\right]^5 \quad P\left(\frac{5s}{6}\right)$$

for the two cases where integer arithmetic is employed. The error for floating-point data remains the same as before. The probability densities for unscaled data, data scaled by 16, and floating-point data are graphed in Figure III-7. Maximum possible errors are increased by one-half count during integer roundoff, and the three maximum errors are 4 1/9, 29 1/9, and 1 2/3 counts.

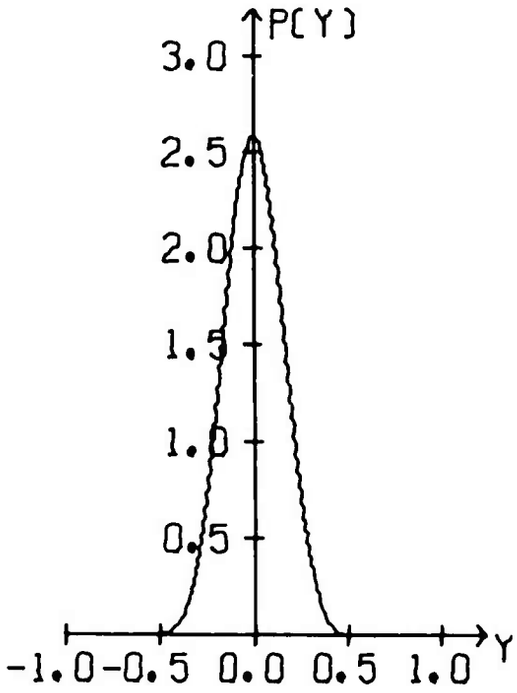
As the various probability densities were computed, standard deviations for the error were calculated. Table III-1 is a summary of the standard deviation and maximum possible error at each stage for the three modes of processing examined.



(a) Unscaled Data



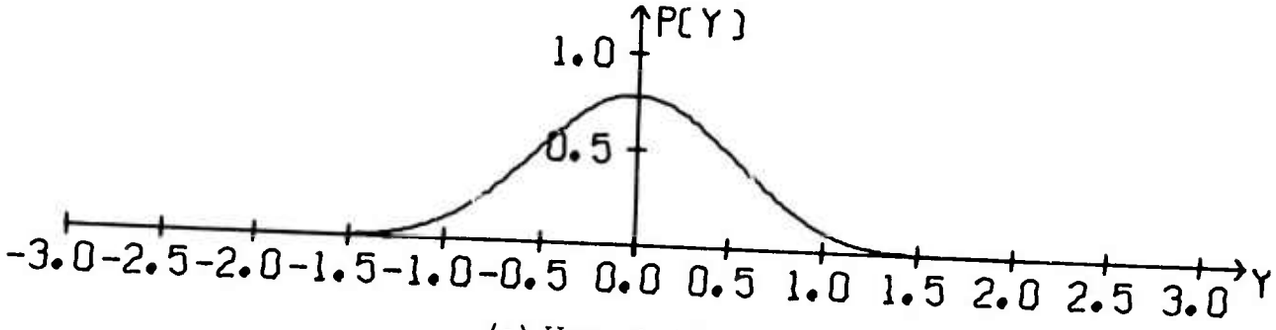
(b) Data Scaled by 16



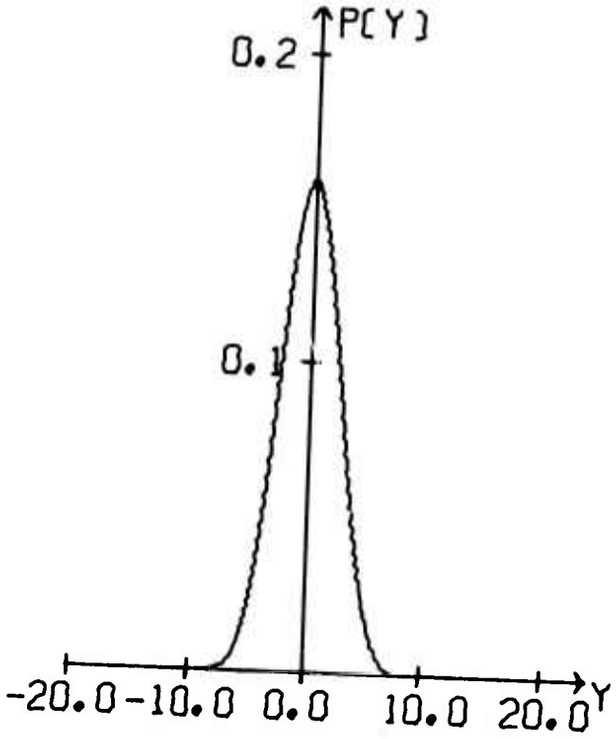
(c) Floating-Point Data

FIGURE III-6

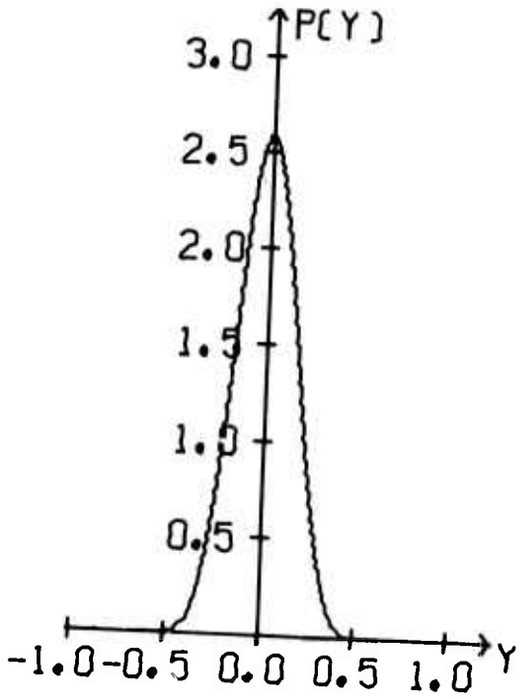
ERROR PROBABILITY DENSITY PER COUNT IN THE TERM $\bar{x}(t) - x_i(t)$ (BEAMSTEER OUTPUT MINUS A SINGLE CHANNEL) BEFORE BEAMSTEER ROUNDOFF



(a) Unscaled Data



(b) Data Scaled by 16



(c) Floating-Point Data

FIGURE III-7

ERROR PROBABILITY DENSITY PER COUNT IN THE TERM $\bar{x}(t) - x_i(t)$ (BEAMSTEER OUTPUT MINUS A SINGLE CHANNEL) AFTER BEAMSTEER ROUNDOFF

TABLE III-1

STANDARD DEVIATION OF ERROR AND MAXIMUM ERROR AT EACH STAGE OF PROCESSING FOR UNSCALED DATA, DATA SCALED BY 16, AND FLOATING-POINT DATA

	Standard Deviation of Error (Counts)		
	Unscaled Data	Data Scaled by 16	Floating-Point Data
Rotated Data	0.319	2.681 (0.168 x 16)	0.167
Prefiltered Data	0.426	2.642 (0.165 x 16)	0.163
Beamsteer Output			
a) Without Beamsteer Roundoff	0.174	1.079 (0.067 x 16)	0.067
b) With Beamsteer Roundoff	0.344	1.119 (0.070 x 16)	0.067
Beamsteer Output Minus Single Channel			
a) Without Beamsteer Roundoff	0.388	2.412 (0.151 x 16)	0.149
b) With Beamsteer Roundoff	0.489	2.430 (0.152 x 16)	0.149
Beamsteer Roundoff	0.297	0.297 (0.019 x 16)	0.000
	Maximum Error (Counts)		
	Unscaled Data	Data Scaled by 16	Floating-Point Data
Rotated Data	0.833	8.333 (0.521 x 16)	0.500
Prefiltered Data	2.167	17.167 (1.073 x 16)	1.000
Beamsteer Output			
a) Without Beamsteer Roundoff	2.167	17.167 (1.073 x 16)	1.000
b) With Beamsteer Roundoff	2.667	17.667 (1.104 x 16)	1.000
Beamsteer Output Minus Single Channel			
a) Without Beamsteer Roundoff	3.611	28.611 (1.788 x 16)	1.667
b) With Beamsteer Roundoff	4.111	29.111 (1.819 x 16)	1.667
Beamsteer Roundoff	0.500	0.500 (0.031 x 16)	0.000

C. DIRECTIONAL ERROR OF THE VECTOR $(\bar{X}-X)$

1. Synopsis of this Subsection

In the next part of this subsection, the geometrical relationships between error in the vector $(\bar{X}-X)$ and the maximum-likelihood constraints on the filter update vector $(A^{new}-A^{old})$ are explored. By separating the error vector into the part E_c due to error before beamsteer roundoff and the part E_b due to the roundoff process occurring at the very end of the beamsteer output formation, it is possible to show that the vector E_c lies within the constraint space and the vector E_b is normal to the constraint space.

In part three, probability distributions for the squared magnitudes $|E_c|^2$, $|E_b|^2$, and $|\bar{X}-X|^2$ are specified.

In part four, the probability distribution is given for the angle α between the vectors E_c and $(\bar{X}-X)$ within the constraint space.

In part five, the probability distribution for the angle θ between the vector $(\bar{X}-X)+E_c$ and $(\bar{X}-X)$ is derived and evaluated. The angle θ is the angle between the vector $(\bar{X}-X)$ and the projection of the vector $(\bar{X}-X)+E_c+E_b$ onto the constraint space.

In the final part of this subsection, a mathematical expression for the probability of the angle between the vector $(\bar{X}-X)+E_c+E_b$ and the vector $(\bar{X}-X)$ is presented. This angle is the angle of error in the vector $(\bar{X}-X)$.

2. Geometrical Relationships Between the Error in $(\bar{X}-X)$ and the Maximum-Likelihood Constraint Conditions

The maximum-likelihood filter set is designed to pass, with unity response at all frequencies, a signal from the desired look direction. Provided that the data channels have been time-shifted according to the beamsteer time delays, this requirement results in the 2N+1 equations

$$\sum_{i=1}^{M} a_i(j) = \delta_{jo} \qquad (j = -N, \ldots, -1, 0, 1, \ldots, N),$$

where δ_{jo} is the Kronecker delta operator. If the filter change vector $(A^{new} - A^{old})$ satisfies the relationships

$$\sum_{i=1}^{M} \left[a_i^{new}(j) - a_i^{old}(j) \right] = 0$$

and the old filter vector satisfies the maximum-likelihood constraints, the new filter vector also satisfies the constraints:

$$\sum_{i=1}^{M} a_i^{new}(j) = \sum_{i=1}^{M} a_i^{old}(j) + \sum_{i=1}^{M} \left[a_i^{new}(j) - a_i^{old}(j) \right] = \delta_{jo} .$$

Henceforth the constraint space will refer to the set of vectors satisfying the restrictions

$$\sum_{i=1}^{M} \left[a_i^{new}(j) - a_i^{old}(j) \right] = 0$$

imposed on the filter change vector for all lags j. These $2N+1$ equations each define a plane of dimension $M-1$ in the M-dimensional subspace corresponding to the j-th lag. If unit vectors U_j (each within the subspace associated with the j-th lag) are defined by the equation

$$U_j = \frac{1}{\sqrt{M}} \begin{bmatrix} \begin{bmatrix} \delta_{j,\,-N} \\ \vdots \\ \delta_{j,\,-N} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \delta_{jo} \\ \vdots \\ \delta_{jo} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \delta_{j\,N} \\ \vdots \\ \delta_{j\,N} \end{bmatrix} \end{bmatrix} \qquad (j = -N, \ldots, -1, 0, 1, \ldots, N),$$

the limitations on the filter change vector may be expressed with vector notation in the $2N+1$ equations

$$U_j^T (A^{new} - A^{old}) = 0 .$$

In the subspace connected with lag j, U_j is a unit vector normal to the constraint plane. The vector $(\bar{X}-X)$ necessarily lies within the constraint space:

$$U_j^T (\bar{X}-X) = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} \left[\bar{x}(t-j) - x_i(t-j) \right]$$

$$= \frac{1}{\sqrt{M}} \left[M\,\bar{x}(t-j) - \sum_{i=1}^{M} x_i(t-j) \right]$$

$$= 0$$

since

$$\bar{x}(t-j) = \frac{1}{M} \sum_{i=1}^{M} x_i(t-j) .$$

In the previous subsection, it was demonstrated that the error in a single component of $(\bar{X}-X)$ is

$$\epsilon_b(t-j) + \frac{1}{M}\left[\sum_{k=1}^{M} \epsilon_k(t-j)\right] - \epsilon_i(t-j) ,$$

where $\epsilon_b(t-j)$ is the beamsteer roundoff error at time $t-j\Delta t$ and where $\epsilon_k(t-j)$ and $\epsilon_i(t-j)$ are the errors in the prefiltered vertical component at time $t-j\Delta t$ for the k-th and i-th channels, respectively. The error vector may be broken down into the portion E_b due to beamsteer roundoff error and the portion E_c due to error in the prefiltered vertical component outputs. Since the beamsteer roundoff error is identical for all channels at time $t-j\Delta t$, the vector E_b is

$$E_b = \sqrt{M} \sum_{j=-N}^{N} \epsilon_b(t-j)\, U_j ,$$

a vector normal to the constraint space. The space of all such possible vectors E_b is spanned by the $2N+1$ orthonormal vectors U_j. It is therefore of dimension $2N+1$.

The vector E_c, on the other hand, satisfies the constraint equations:

$$U_j^T E_c = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} \left\{ \left[\frac{1}{M} \sum_{k=1}^{M} \epsilon_k(t-j) \right] - \epsilon_i(t-j) \right\}$$

$$= \frac{1}{\sqrt{M}} \left[\sum_{k=1}^{M} \epsilon_k(t-j) - \sum_{i=1}^{M} \epsilon_i(t-j) \right]$$

$$= 0$$

And so it must lie entirely within the constraint space. In the subspace coinciding with time $t-j\Delta t$, the vectors

$$U_j = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \qquad W_{1j} = \frac{1}{\sqrt{M(M-1)}} \begin{bmatrix} M-1 \\ -1 \\ \vdots \\ -1 \end{bmatrix},$$

$$W_{2j} = \frac{1}{\sqrt{(M-1)(M-2)}} \begin{bmatrix} 0 \\ M-2 \\ -1 \\ \vdots \\ -1 \end{bmatrix}, \qquad W_{3j} = \frac{1}{\sqrt{(M-2)(M-3)}} \begin{bmatrix} 0 \\ 0 \\ M-3 \\ -1 \\ \vdots \\ -1 \end{bmatrix}, \ldots,$$

$$W_{M-2,j} = \frac{1}{\sqrt{3 \cdot 2}} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 2 \\ -1 \\ -1 \end{bmatrix}, \text{ and } W_{M-1,j} = \frac{1}{\sqrt{2 \cdot 1}} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

form an orthonormal basis. The inner product between E_c and W_{ij} is

$$E_c^T W_{ij} = \frac{\left[\sum_{p=i+1}^{M} \epsilon_p(t-j)\right] - (M-i)\,\epsilon_i(t-j)}{\sqrt{(M+1-i)\,(M-i)}}$$

$$= \frac{(M+1-i)\left\{\left[\frac{1}{M}\sum_{k=1}^{M} \epsilon_k(t-j)\right] - \epsilon_i(t-j)\right\} + \sum_{p=1}^{i-1}\left\{\left[\frac{1}{M}\sum_{k=1}^{M} \epsilon_k(t-j)\right] - \epsilon_p(t-j)\right\}}{\sqrt{(M+1-i)\,(M-i)}}$$

This inner product is not identically zero, so that the vectors E_c fill the space

spanned by the orthonormal vectors W_{ij} ($i = 1, 2, \ldots, M-1$; $j = -N, \ldots, -1, 0, 1, \ldots, N$). The space of all possible vectors E_c, therefore, is of dimension $(M-1)(2N+1)$.

3. Probability Distributions for the Squared Magnitudes $|E_c|^2$, $|E_b|^2$, and $|\bar{X}-X|^2$

Let the vinculum operator $\bar{Q}$ for any function Q of the $M(2N+1)$ variables $\epsilon_i(t-j)$ denote the mean of the function Q over all possible values of $\epsilon_i(t-j)$:

$$\bar{Q} = \left[\prod_{i=1}^{M} \prod_{j=-N}^{N} \int_{\epsilon_i(t-j)=-\infty}^{\epsilon_i(t-j)=\infty} \right] Q \left\{ \prod_{i=1}^{M} \prod_{j=-N}^{N} p\left[\epsilon_i(t-j)\right] d\,\epsilon_i(t-j) \right\}.$$

The mean μ_c of the squared magnitude $|E_c|^2$ of the error vector E_c within the constraint space is

$$\mu_c = \overline{\sum_{j=-N}^{N} \sum_{i=1}^{M} \left[\epsilon_i(t-j) - \frac{1}{M} \sum_{k=1}^{M} \epsilon_k(t-j) \right]^2}$$

$$= \overline{\sum_{j=-N}^{N} \sum_{i=1}^{M} \left[\epsilon_i^2(t-j) - \frac{2\,\epsilon_i(t-j)}{M} \sum_{k=1}^{M} \epsilon_k(t-j) + \frac{1}{M^2} \sum_{k=1}^{M} \sum_{p=1}^{M} \epsilon_k(t-j)\,\epsilon_p(t-j) \right]}$$

$$= \overline{\sum_{j=-N}^{N} \left[\sum_{i=1}^{M} \epsilon_i^2(t-j) - \frac{2}{M} \sum_{i=1}^{M} \sum_{k=1}^{M} \epsilon_i(t-j)\,\epsilon_k(t-j) + \frac{1}{M} \sum_{k=1}^{M} \sum_{p=1}^{M} \epsilon_k(t-j)\,\epsilon_p(t-j) \right]}$$

$$= \overline{\sum_{j=-N}^{N} \left[\sum_{i=1}^{M} \epsilon_i^2(t-j) - \frac{1}{M} \sum_{i=1}^{M} \sum_{k=1}^{M} \epsilon_i(t-j)\,\epsilon_k(t-j) \right]}$$

$$= \sum_{j=-N}^{N} \left\{ \left[\frac{M-1}{M} \sum_{i=1}^{M} \epsilon_i^2(t-j) \right] - \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ i \neq k}}^{M} \epsilon_i(t-j)\, \epsilon_k(t-j) \right\}$$

$$= \frac{M-1}{M} \sum_{j=-N}^{N} \sum_{i=1}^{M} \overline{\epsilon_i^2(t-j)}$$

$$= (M-1)(2N+1)\, \overline{\epsilon_i^2(t-j)}$$

provided that the zero-mean random variables $\epsilon_i(t-j)$ are mutually independent and identically distributed.

The variance σ_c^2 of $|E_c|^2$ is

$$\sigma_c^2 = \overline{\left[\left\{ \sum_{j=-N}^{N} \sum_{i=1}^{M} \left[\epsilon_i(t-j) - \frac{1}{M} \sum_{k=1}^{M} \epsilon_k(t-j) \right]^2 \right\} - \mu_c \right]^2}$$

$$= \overline{\left\{ \left[\sum_{j=-N}^{N} \left\{ \left[\frac{M-1}{M} \sum_{i=1}^{M} \epsilon_i^2(t-j) \right] - \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ i \neq k}}^{M} \epsilon_i(t-j)\, \epsilon_k(t-j) \right\} \right] - \mu_c \right\}^2}$$

$$= \overline{\left[\sum_{j=-N}^{N} \left\{ \frac{M-1}{M} \sum_{i=1}^{M} \left[\epsilon_i^2(t-j) - \overline{\epsilon_i^2(t-j)} \right] - \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ i \neq k}}^{M} \epsilon_i(t-j)\, \epsilon_k(t-j) \right\} \right]^2}$$

$$= \left(\frac{M-1}{M} \right)^2 \overline{\sum_{j=-N}^{N} \sum_{q=-N}^{N} \sum_{i=1}^{M} \sum_{k=1}^{M} \left[\epsilon_i^2(t-j) - \overline{\epsilon_i^2(t-j)} \right] \left[\epsilon_k^2(t-q) - \overline{\epsilon_k^2(t-q)} \right]}$$

$$- \frac{2(M-1)}{M^2} \sum_{j=-N}^{N} \sum_{q=-N}^{N} \sum_{i=1}^{M} \sum_{k=1}^{M} \sum_{\substack{p=1 \\ k\neq p}}^{M} \overline{\left[\epsilon_i^2(t-j) - \overline{\epsilon_i^2(t-j)}\right] \epsilon_k(t-q)\ \epsilon_p(t-q)}$$

$$+ \frac{1}{M^2} \sum_{j=-N}^{N} \sum_{q=-N}^{N} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ i\neq k}}^{M} \sum_{\substack{\ell=1 \\ \ell\neq p}}^{M} \sum_{p=1}^{M} \overline{\epsilon_i(t-j)\, \epsilon_k(t-j)\, \epsilon_\ell(t-q)\, \epsilon_p(t-q)}$$

$$= \left(\frac{M-1}{M}\right)^2 \sum_{j=-N}^{N} \sum_{i=1}^{M} \overline{\left[\epsilon_i^2(t-j) - \overline{\epsilon_i^2(t-j)}\right]^2}$$

$$+ \frac{1}{M^2} \sum_{j=-N}^{N} \sum_{\substack{i=1 \\ i\neq k}}^{M} \sum_{k=1}^{M} \left[\overline{\epsilon_i(t-j)\, \epsilon_k(t-j)\, \epsilon_i(t-j)\, \epsilon_k(t-j)} + \overline{\epsilon_i(t-j)\, \epsilon_k(t-j)\, \epsilon_k(t-j)\, \epsilon_i(t-j)}\right]$$

$$= \left(\frac{M-1}{M}\right)^2 \sum_{j=-N}^{N} \sum_{i=1}^{M} \overline{\left[\epsilon_i^2(t-j) - \overline{\epsilon_i^2(t-j)}\right]^2}$$

$$+ \frac{2}{M^2} \sum_{j=-N}^{N} \sum_{\substack{i=1 \\ i\neq k}}^{M} \sum_{k=1}^{M} \overline{\epsilon_i^2(t-j)}\ \overline{\epsilon_k^2(t-j)}$$

$$= \frac{(M-1)(2N+1)}{M} \left\{ (M-1) \overline{\left[\epsilon_i^2(t-j) - \overline{\epsilon_i^2(t-j)}\right]^2} + 2\left[\overline{\epsilon_i^2(t-j)}\right]^2 \right\}$$

under the same assumptions as before. The quantity inside the braces is (M-1) times the variance of $\epsilon_i^2(t-j)$ plus twice the squared mean of $\epsilon_i^2(t-j)$. If the unsquared random variable $\epsilon_i(t-j)$ were normally distributed, the following equations would be valid:

$$\overline{\left[\epsilon_i^2(t-j) - \overline{\epsilon_i^2(t-j)}\right]^2} = 2\left[\overline{\epsilon_i^2(t-j)}\right]^2 ;$$

$$\sigma_c^2 = (M-1)(2N+1)\left[\overline{\epsilon_i^2(t-j)} - \overline{\epsilon_i^2(t-j)}\right]^2 ;$$

$$\sigma_c^2 = 2(M-1)(2N+1)\left[\overline{\epsilon_i^2(t-j)}\right]^2 .$$

In that event, these equations could have been deduced by trivial manipulations of the χ^2-distribution for $(M-1)(2N+1)$ dimensions. Unfortunately, the pre-filtered vertical-component output error $\epsilon_i(t-j)$ is not normally distributed, as can be seen from Table III-2, which lists the variance of $\epsilon_i^2(t-j)$ together with the ratio of its variance to its squared mean for unscaled data, data scaled by 16, and floating-point data. The ratio in the righthand column is sufficiently different from two that the non-Gaussian character of $\epsilon_i(t-j)$ must be taken into account.

The random variable $|E_c|^2$, on the other hand, is the sum of the $2N+1$ (i.e., 31) independent, identically-distributed random variables

$$\sum_{i=1}^{M}\left[\epsilon_i(t-j) - \frac{1}{M}\sum_{k=1}^{M}\epsilon_k(t-j)\right]^2 .$$

The Central Limit Theorem will be assumed valid, so that the probability distribution of $|E_c|^2$ can be approximated by a normal distribution with mean μ_c and variance σ_c^2. Table III-3 gives the mean, standard deviation, and variance of $|E_c|^2$ for the three cases being considered. Figure III-8 is a plot of the corresponding probability density functions for all three cases.

The random variable $|E_b|^2$, similarly, is the sum of the $2N+1$ independent, identically-distributed random variables $M\epsilon_b^2(t-j)$. In the two cases involving integer arithmetic, $|E_b|^2$ has the same probability distribution. With six channels, the probability mass function for the unsquared random variable $\epsilon_b(t-j)$ is

TABLE III-2

VARIANCE AND RATIO OF VARIANCE TO SQUARED MEAN FOR THE RANDOM VARIABLE $\epsilon_i^2(t-j)$ (SQUARED PREFILTERED VERTICAL-COMPONENT OUTPUT ERROR)

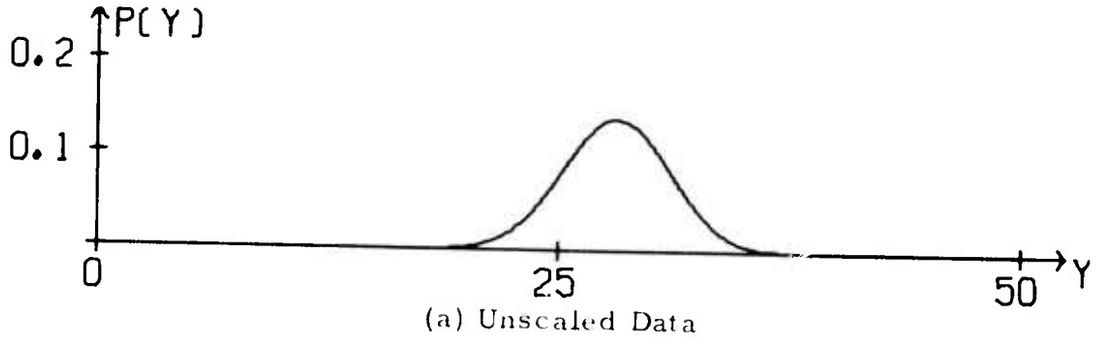
	Variance (Counts^4)	Variance/(Mean)^2
Unscaled Data	4.991×10^{-2}	1.522
Data Scaled by 16	8.003×10 $(1.221 \times 10^{-3} \times 65536)$	1.643 1.643
Floating-Point Data	1.156×10^{-3}	1.627

TABLE III-3

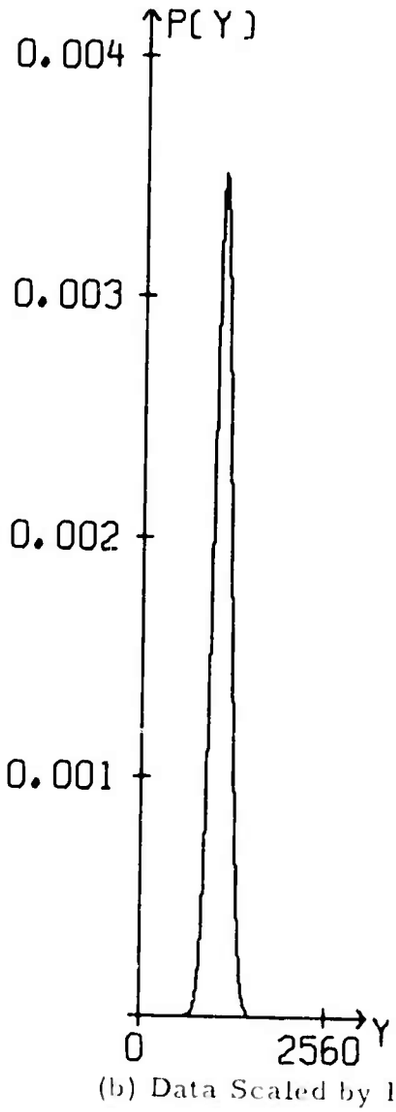
MEAN, STANDARD DEVIATION, AND VARIANCE OF $|E_c|^2$

(SQUARED MAGNITUDE OF ERROR VECTOR PROJECTED ONTO CONSTRAINT SPACE)

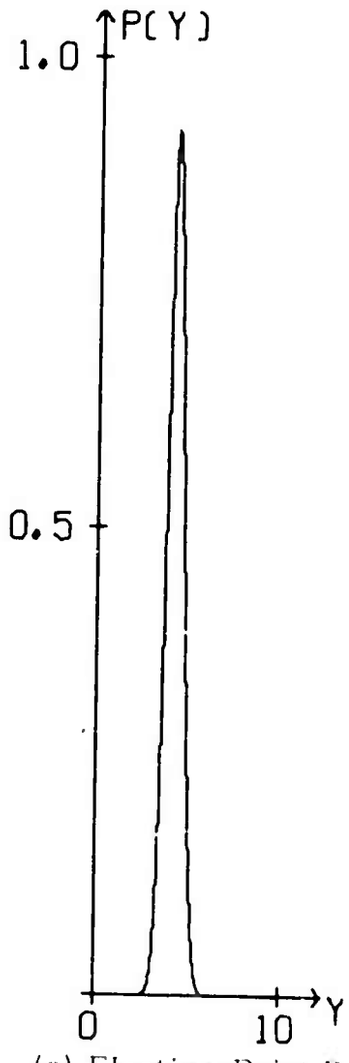
	Mean (Counts^2)	Standard Deviation (Counts^2)	Variance (Counts^4)
Unscaled Data	28.071	2.853	8.142
Data Scaled by 16	1081.952 (4.226 x 256)	113.380 (0.443 x 256)	12855.064 (0.196 x 65536)
Floating-Point Data	4.133	0.431	0.186



(a) Unscaled Data



(b) Data Scaled by 16



(c) Floating-Point Data

FIGURE III-8

PROBABILITY DENSITY PER COUNT SQUARED FOR THE RANDOM VARIABLE $|E_c|^2$
(SQUARED MAGNITUDE OF ERROR VECTOR PROJECTED ONTO CONSTRAINT SPACE)

$$p(y) = 1/12 \qquad (y = -1/2,\ 1/2)$$

$$p(y) = 1/6 \qquad (y = -1/3,\ -1/6,\ 0,\ 1/6,\ 1/3)\ ,$$

so that the probability mass function for $M\epsilon_b^2(t\text{-}j)$ is

$$p(y) = 1/6 \qquad (y = 0,\ 1\ 1/2)$$

$$p(y) = 1/3 \qquad (y = 1/6,\ 2/3)\ .$$

The probability mass function for $|E_b|^2$ is obtained by convolving this function with itself $2N+1$ times. Figure III-9 pictures the resulting probability mass function. Probability is concentrated at discrete points between 0 and 46 1/2 at intervals of 1/6. The vertical lines give the probability for each discrete point. The mean, standard deviation, and variance are 16.361 counts^2, 2.805 counts^2, and 7.870 counts^4, respectively. Note that the plot is slightly skewed: the mode point (16.167 counts^2) is slightly below the mean, and the probability envelope falls off somewhat less rapidly on the right. When floating point arithmetic is used, the beamsteer roundoff error is assumed to be negligible, so that $|E_b|^2 = 0$.

In the case of the random variable $|\bar{X}\text{-}X|^2$, the probability distribution may be measured. A four-hour noise sample from day 238 of 1970 was used for this purpose. The quantity $|\bar{X}\text{-}X|^2$ was computed every 31 points using integer data scaled by a factor of 16. The resulting values were divided by 256 and sorted by magnitude to produce the cumulative distribution function and histogram of Figure III-10. In the histogram, a bin width of 1000 counts^2 was utilized. The vertical axis indicates how many times values of $|\bar{X}\text{-}X|^2$ occurred within the limits of a particular bin. The probability distribution of Figure III-10 applies both to the case of unscaled data and floating-point data. To obtain the equivalent distribution for data scaled by 16, the horizontal axis needs to be multiplied by 256.

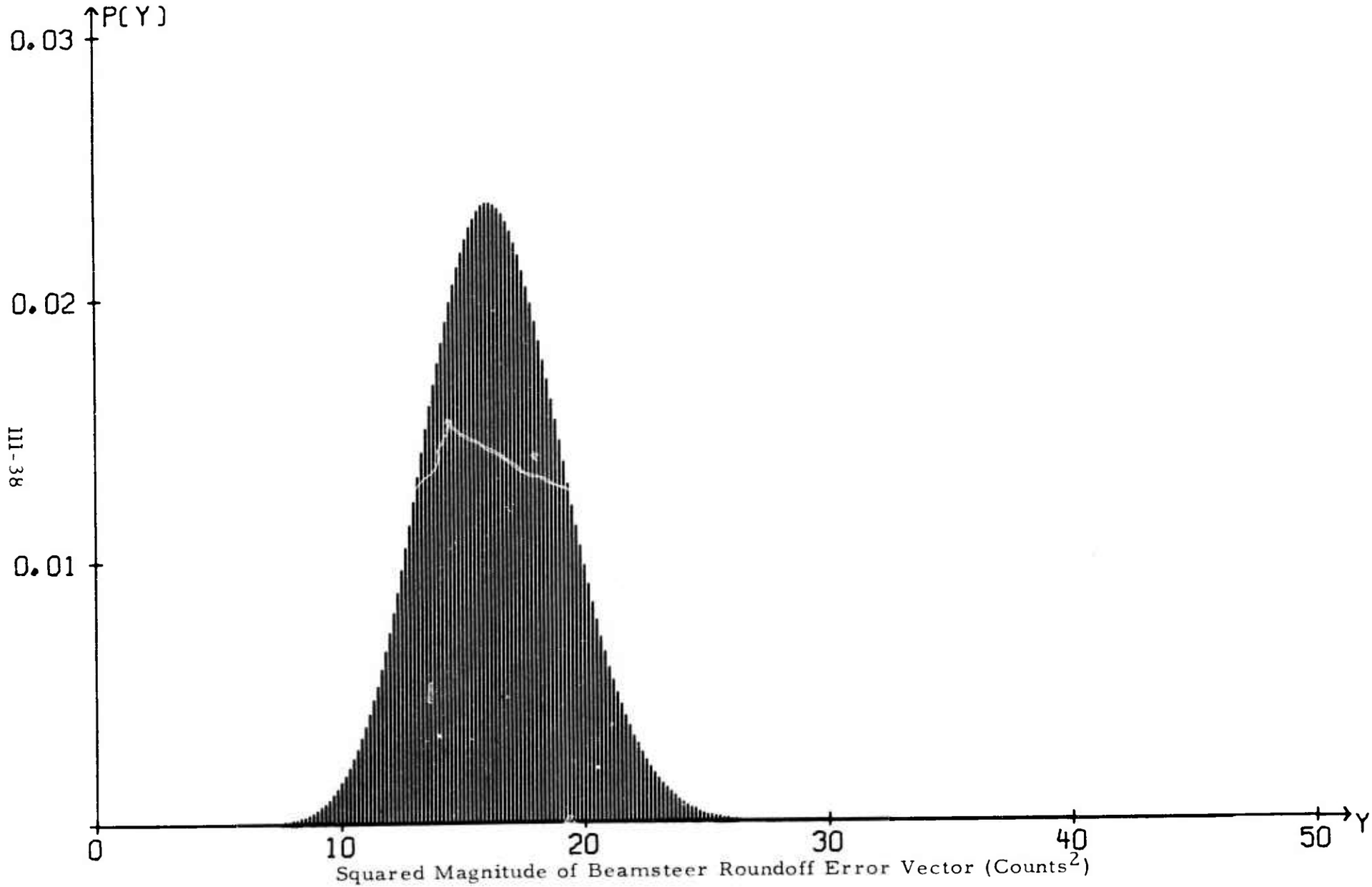
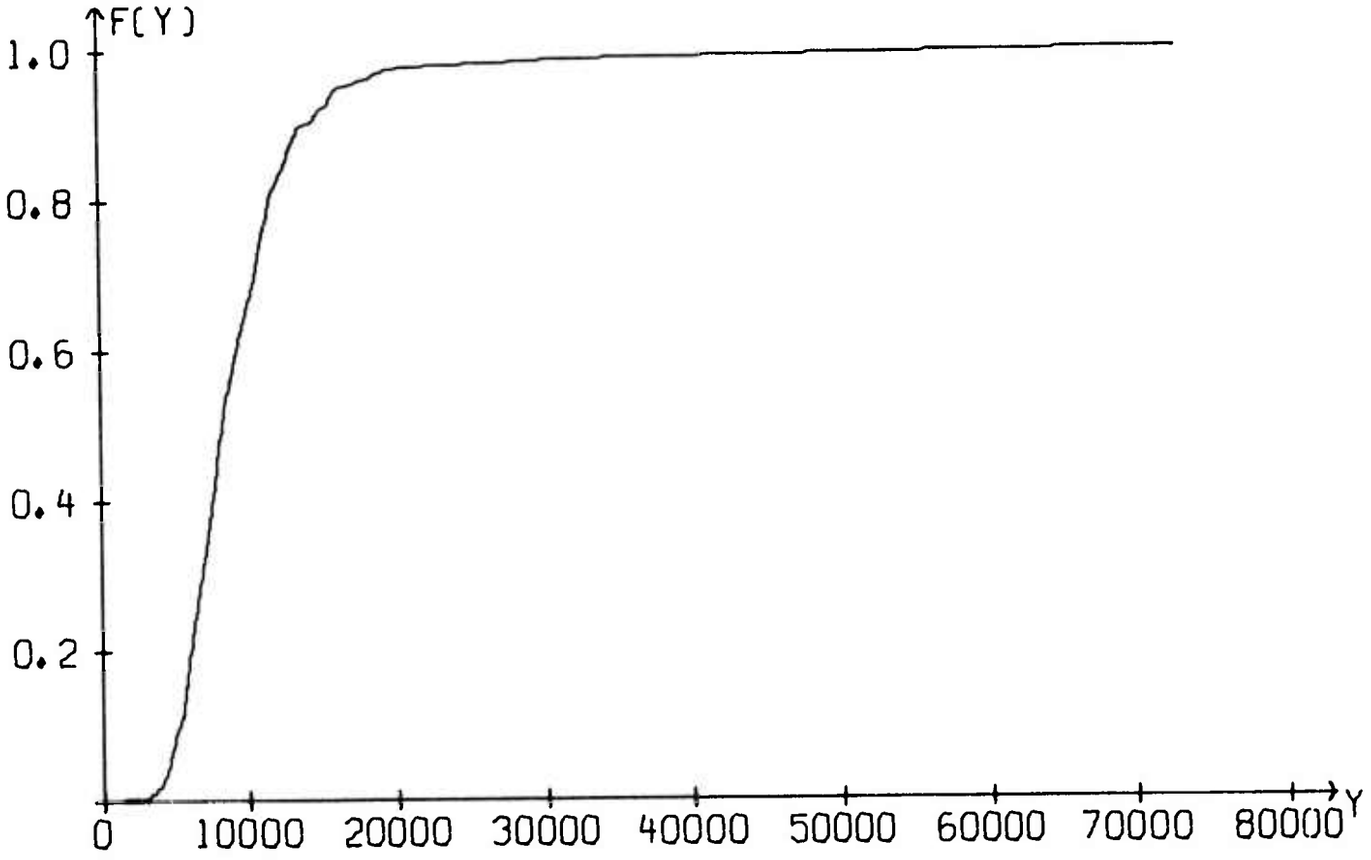


FIGURE III-9

PROBABILITY MASS FUNCTION FOR THE RANDOM VARIABLE $|E_b|^2$



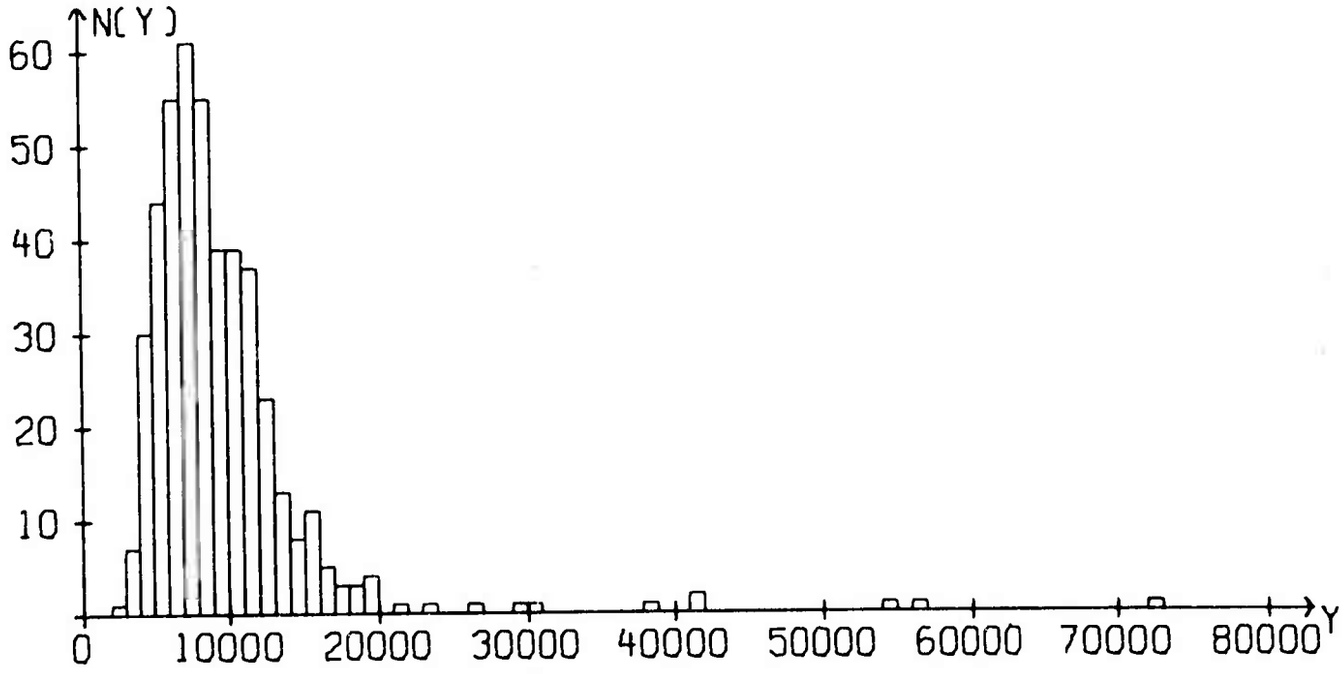


FIGURE III-10

CUMULATIVE DISTRIBUTION FUNCTION AND HISTOGRAM FOR THE RANDOM VARIABLE $|\overline{X}-X|^2$ (FOR A NOISE SAMPLE COVERING THE INTERVAL 0757-1150 ON DAY 238 OF 1970)

4. Probability Distribution For The Angle α Between The Vectors E_c and $(\bar{X}-X)$

Within the constraint space, the direction of $(\bar{X}-X)$ is presumed to be independent of the direction of E_c. This assumption may be satisfied in a number of ways. For example, if the multidimensional probability density function for either E_c or $(\bar{X}-X)$ were dependent solely on the magnitude of the vector, this assumption would be satisfied. Such would be the case if the components of either vector were mutually independent and normally distributed. More realistically, if the preferred orientations of one vector were weak and not too strongly related to the preferred orientations of the other vector, the assumption would be a good approximation.

Under the assumption just described, the probability that the angle between E_c and $(\bar{X}-X)$ is less than the angle α is equal to proportion of a sphere subtended by a cone of angle α in a space of $n = (M-1)(2N+1)$ dimensions, where n is the dimension of the constraint space ($n = 155$ for 6 channels and 31 filter points per channel). The axis of the cone may lie along any direction in the constraint space.

To find the proportion of the sphere subtended by a cone of angle α, the polar coordinate transformation

$$
\begin{aligned}
z_1^n &= \rho \cos \theta_{n-1} \\
z_2^n &= \rho \sin \theta_{n-1} \cos \theta_{n-2} \\
&\vdots \\
z_{n-1}^n &= \rho \sin \theta_{n-1} \sin \theta_{n-2} \cdots \sin \theta_2 \cos \theta_1 \\
z_n^n &= \rho \sin \theta_{n-1} \sin \theta_{n-2} \cdots \sin \theta_2 \sin \theta_1
\end{aligned}
$$

is made. The superscript n denotes the dimension of the space in which the

transformation is made. The Jacobian

$$\frac{\partial\left(z_1^n,\ z_2^n,\ \ldots,\ z_n^n\right)}{\partial\left(\rho,\ \theta_{n-1},\ \ldots,\ \theta_1\right)},$$

expanded in terms of minors relative to the leftmost column of the determinant, is

$$\frac{\partial\left(z_1^n,\ z_2^n,\ \ldots,\ z_n^n\right)}{\partial\left(\rho,\ \theta_{n-1},\ \ldots,\ \theta_1\right)}$$

$$= \frac{\partial z_1^n}{\partial\rho}\ \frac{\partial\left(z_2^n,\ z_3^n,\ \ldots,\ z_n^n\right)}{\partial\left(\theta_{n-1},\ \theta_{n-2},\ \ldots,\ \theta_1\right)} - \frac{\partial z_1^n}{\partial\theta_{n-1}}\ \frac{\partial\left(z_2^n,\ z_3^n,\ \ldots,\ z_n^n\right)}{\partial\left(\rho,\theta_{n-2},\ \ldots,\ \theta_1\right)}$$

since $\partial z_1^n \Big/ \partial\theta_i = 0$ for all values of i from 1 to n-2. The following partial derivatives are easily evaluated:

$$\frac{\partial z_1^n}{\partial\rho} = \cos\theta_{n-1}$$

$$\frac{\partial z_1^n}{\partial\theta_{n-1}} = -\rho\sin\theta_{n-1}$$

The corresponding polar coordinate transformation in the next lower dimension is

$$z_1^{n-1} = \rho\cos\theta_{n-2}$$

$$z_2^{n-1} = \rho\sin\theta_{n-2}\cos\theta_{n-3}$$

$$\vdots$$

$$z_{n-2}^{n-1} = \rho\sin\theta_{n-2}\sin\theta_{n-3}\ \ldots\ \sin\theta_2\cos\theta_1$$

$$z_{n-1}^{n-1} = \rho\sin\theta_{n-2}\sin\theta_{n-3}\ \ldots\ \sin\theta_2\sin\theta_1$$

Note that, for all values of i from 1 to n-1,

$$z_{i+1}^{n} = \sin\theta_{n-1}\, z_i^{n-1}$$

so that

$$\frac{\partial z_{i+1}^{n}}{\partial \rho} = \sin\theta_{n-1}\, \frac{\partial z_i^{n-1}}{\partial \rho} \quad ,$$

$$\frac{\partial z_{i+1}^{n}}{\partial \theta_j} = \sin\theta_{n-1}\, \frac{\partial z_i^{n-1}}{\partial \theta_j}$$

for all values of j from 1 to n-2, and

$$\frac{\partial z_{i+1}^{n}}{\partial \theta_{n-1}} = \cos\theta_{n-1}\, z_i^{n-1} = \rho\cos\theta_{n-1}\, \frac{\partial z_i^{n-1}}{\partial \rho}$$

since $\rho\, \dfrac{\partial z_i^{n-1}}{\partial \rho} = z_i^{n-1}$. Therefore,

$$\frac{\partial\left(z_1^n, z_2^n, \ldots, z_n^n\right)}{\partial\left(\rho, \theta_{n-1}, \ldots, \theta_1\right)} = \rho\cos^2\theta_{n-1}\sin^{n-2}\theta_{n-1}\frac{\partial\left(z_1^{n-1}, z_2^{n-1}, \ldots, z_{n-1}^{n-1}\right)}{\partial\left(\rho, \theta_{n-2}, \ldots, \theta_1\right)}$$

$$+ \rho\sin^2\theta_{n-1}\sin^{n-2}\theta_{n-1}\frac{\partial\left(z_1^{n-1}, z_2^{n-1}, \ldots, z_{n-1}^{n-1}\right)}{\partial\left(\rho, \theta_{n-2}, \ldots, \theta_1\right)}$$

$$= \rho\sin^{n-2}\theta_{n-1}\frac{\partial\left(z_1^{n-1}, z_2^{n-1}, \ldots, z_{n-1}^{n-1}\right)}{\partial\left(\rho, \theta_{n-2}, \ldots, \theta_1\right)}$$

The Jacobian in n dimensions can be expressed as the product of $\rho\sin^{n-2}\theta_{n-1}$ and the corresponding Jacobian in the next lower dimension. For the case of two dimensions,

$$z_1^2 = \rho\cos\theta_1$$

$$z_2^2 = \rho\sin\theta_1$$

and the Jacobian for two dimensions is

$$\frac{\partial\left(z_1^2,\ z_2^2\right)}{\partial\left(\rho,\theta_1\right)} = \begin{vmatrix} \dfrac{\partial z_1^2}{\partial \rho} & \dfrac{\partial z_2^2}{\partial \rho} \\ \dfrac{\partial z_1^2}{\partial \theta_1} & \dfrac{\partial z_2^2}{\partial \theta_1} \end{vmatrix}$$

$$= \begin{vmatrix} \cos\theta_1 & \sin\theta_1 \\ -\rho\sin\theta_1 & \rho\cos\theta_1 \end{vmatrix}$$

$$= \rho\ (\cos^2\theta_1 + \sin^2\theta_1) = \rho$$

By the appropriate inductive reasoning,

$$\frac{\partial\left(z_1^n,\ z_2^n,\ \ldots,\ z_n^n\right)}{\partial\left(\rho,\theta_{n-1},\ \ldots,\theta_1\right)} = \rho^{n-1}\sin^{n-2}\theta_{n-1}\sin^{n-3}\theta_{n-2}\cdots\sin^1\theta_2 .$$

The volume $V(\alpha, n)$ inside a sphere of radius r and inside a cone of angle in an n-dimensional space is

$$V(\alpha,n) = 2^{n-1}\int_0^{\alpha}\sin^{n-2}\theta_{n-1}\,d\theta_{n-1}\left[\prod_{j=1}^{n-2}\int_0^{\pi/2}\sin^{j-1}\theta_j\,d\theta_j\right]\int_0^r\rho^{n-1}\,d\rho .$$

The surface area $S(\alpha, n)$ on the surface of a sphere of radius r and lying inside a cone of angle α is $\partial V(\alpha, n)/\partial r$, so that the proportion of the spherical surface subtended by a cone of angle α is

$$\frac{S(\alpha, n)}{S(\pi, n)} = \frac{\int_0^{\alpha} \sin^{n-2} \theta_{n-1} \, d\theta_{n-1}}{\int_0^{\pi} \sin^{n-2} \theta_{n-1} \, d\theta_{n-1}}$$

This expression is the probability $P(\alpha, n)$ that the angle between two vectors from the center of the sphere to arbitrary points on the surface of the sphere is less than or equal to the angle α (when the probability density at all points of the surface is uniform).

The integrals may be evaluated in closed form using the two formulas

$$\int \sin^{2k} x\,dx = -\cos x \sum_{i=1}^{k} \left\{ \frac{\prod_{j=1}^{i-1} \left[1 + 2(k-j)\right]}{\prod_{j=1}^{i} \left[2(k+1-j)\right]} \right\} \sin^{1+2(k-i)} x$$

$$+ \left\{ \frac{\prod_{j=1}^{k} \left[1 + 2(k-j)\right]}{\prod_{j=1}^{k} \left[2(k+1-j)\right]} \right\} x$$

$$\int \sin^{2k+1} x\,dx = -\cos x \sum_{i=1}^{k+1} \left\{ \frac{\prod_{j=1}^{i-1} \left[2(k+1-j)\right]}{\prod_{j=1}^{i} \left[1 + 2(k+1-j)\right]} \right\} \sin^{2(k+1-i)} x$$

which are valid for all non-negative integers k. When the upper index in a summation sign or product sign is smaller than the corresponding lower index, the results are 0 or 1, respectively. These formulas were obtained by mathematical induction using the formula

$$\int \sin^{k} x\,dx = -\frac{\cos x \sin^{k-1} x}{k} + \frac{(k-1)}{k} \int \sin^{k-2} x\,dx,$$

which can be derived using integration by parts. $P(\alpha, n)$ is a cumulative distribution function for α, so that the probability density function is

$$\frac{\partial P(\alpha, n)}{\partial \alpha} = \frac{\sin^{n-2}\alpha}{\int_0^{\pi} \sin^{n-2}\theta_{n-1}\, d\theta_{n-1}}$$

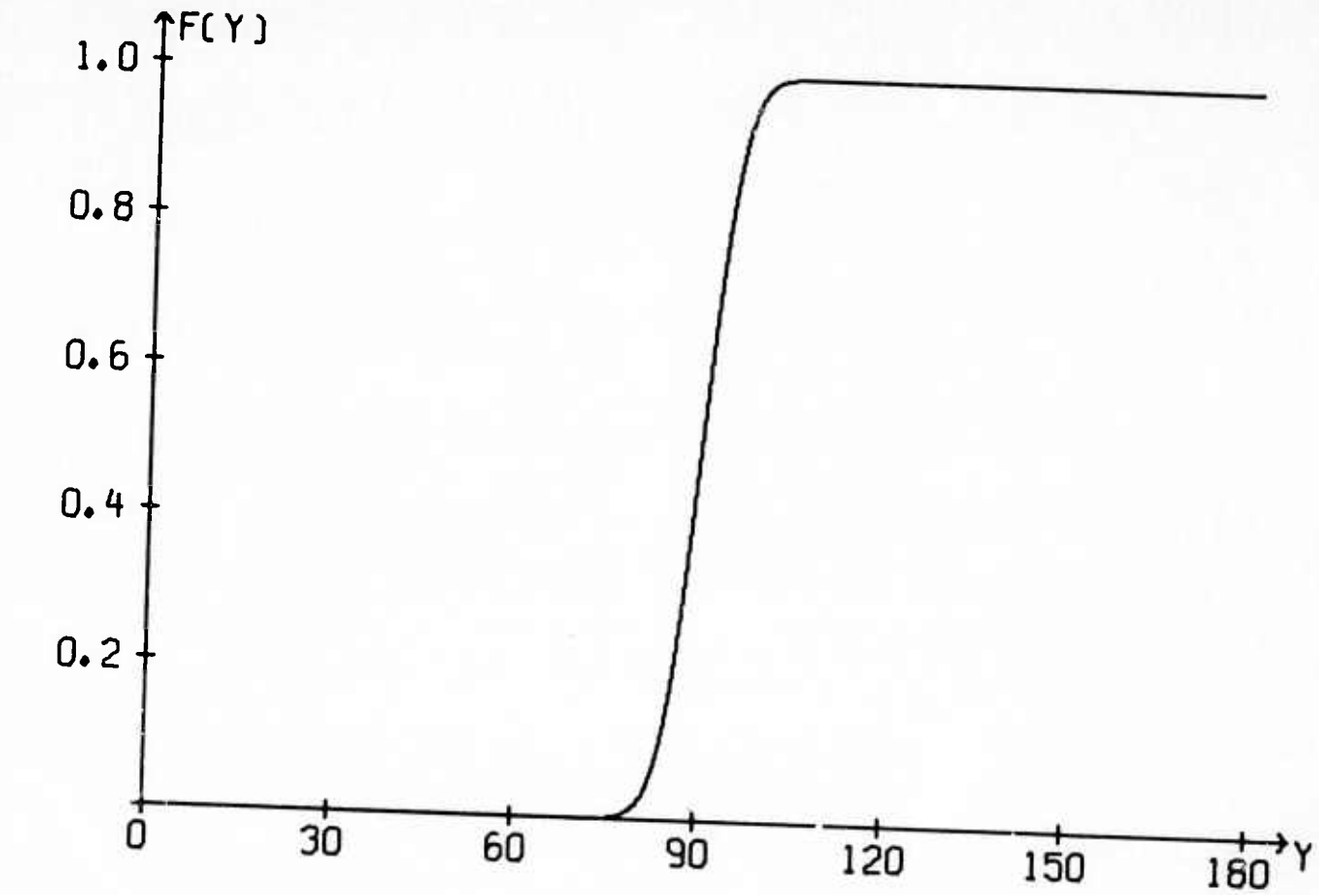
Figure III-11 diagrams the cumulative distribution function and probability density per degree for the angle α between E_c and $(\bar{X}-X)$ in a 155-dimensional constraint space. Note that the probability density function in strongly peaked near $\alpha = 90^{\circ}$. As the dimension of the constraint space increases, the cumulative distribution function approaches a step function jumping from 0 to 1 at 90°. Under the same circumstances, consequently, the probability density function approaches a delta function centered about 90°.

5. Probability Distribution For the Angle θ Between the Vectors $(\bar{X}-X) + E_c$ and $(\bar{X}-X)$

It is possible to construct a two-dimensional plane passing through the origin, any specific vector $(\bar{X}-X)$, and any specific vector E_c. The resulting plane must lie within the constraint space since the origin and the vectors $(\bar{X}-X)$ and E_c lie within the constraint space. Figure III-12 illustrates such a plane. In the figure, $(\bar{X}-X)$ and E_c are as defined previously in this section: α is the angle between the vectors E_c and $(\bar{X}-X)$, and θ is the angle between the vectors $(\bar{X}-X)$ and $(\bar{X}-X) + E_c$.

If E_c is not a scalar multiple of $(\bar{X}-X)$, the angles α and θ are between 0 and π and the sine law holds:

$$\frac{\sin(\alpha-\theta)}{\sin\theta} = \frac{|\bar{X}-X|}{|E_c|} \quad .$$



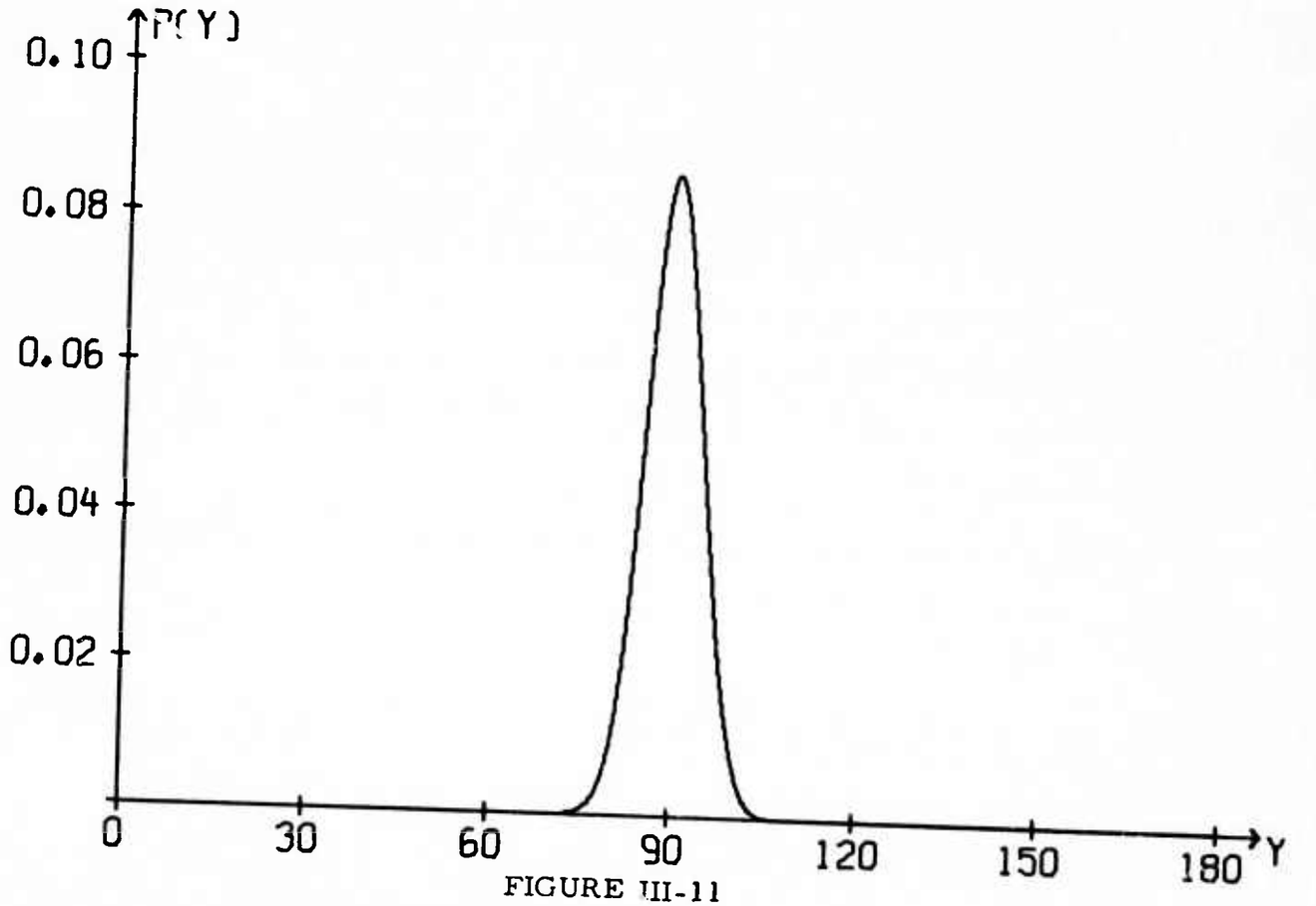


FIGURE III-11

CUMULATIVE DISTRIBUTION FUNCTION AND PROBABILITY DENSITY PER DEGREE FOR THE ANGLE α BETWEEN $(\overline{X}\text{-}X)$ AND E_c IN A 155-DIMENSIONAL CONSTRAINT SPACE

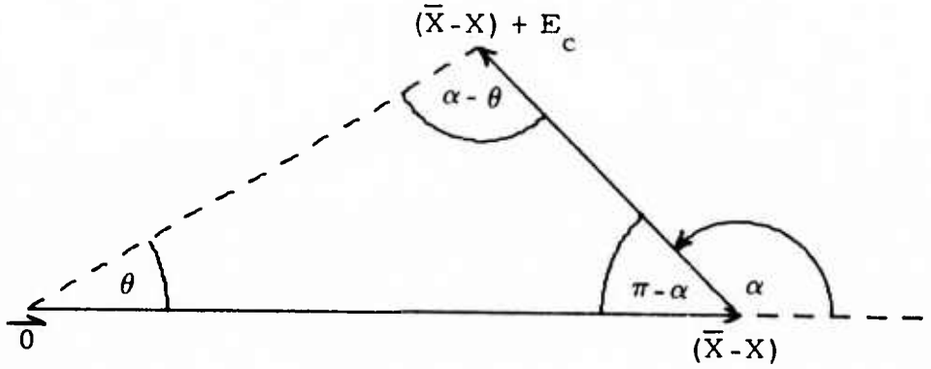


FIGURE III-12

TWO-DIMENSIONAL PLANE PASSING THROUGH THE ORIGIN, $(\bar{X}-X)$, AND E_c

Division by $\sin\alpha$ yields the modified equation

$$\cot\theta - \cot\alpha = \frac{|\bar{X}-X|}{|E_c|\sin\alpha} \quad .$$

Since the cotangent is a monotonically decreasing function over the range 0 to π, $\theta \leq \theta_o$ if and only if $\cot\theta_o \leq \cot\theta$. Thus $\theta \leq \theta_o$ if and only if

$$\cot\theta_o - \cot\alpha \leq \frac{|\bar{X}-X|}{|E_c|\sin\alpha} \quad .$$

This condition is automatically satisfied if $\alpha \leq \theta_o$, since $\cot\theta_o - \cot\alpha \leq 0$ when $\alpha \leq \theta_o$. Hence $\theta \leq \theta_o$ if

$$\alpha \leq \theta_o$$

or if

$$|E_c| \leq \frac{|\bar{X}-X| \sin\theta_o}{\sin(\alpha-\theta_o)}$$

when $\alpha > \theta_o$. It is not necessary to consider the case where E_c is a scalar multiple of $(\bar{X}-X)$, for α must then be either 0 or π: the probability that $\alpha \leq 0$ or $\alpha \geq \pi$ is zero.

Since all quantities in the inequality for $\alpha > \theta_o$ are non-negative, both sides of the inequality may be squared:

$$|E_c|^2 \leq \frac{|X-X|^2 \sin^2\theta_o}{\sin^2(\alpha-\theta_o)} \qquad (\alpha > \theta_o).$$

With a few more minor manipulations, the probability that $\theta \leq \theta_o$ when $\alpha > \theta_o$ becomes the probability that

$$\frac{|E_c|^2 - \mu_c}{\sigma_c} \leq \frac{\mu_c}{\sigma_c}\left[\frac{|X-X|^2 \sin^2\theta_o}{\mu_c \sin^2(\alpha-\theta_o)} - 1\right] ,$$

where μ_c and σ_c are the mean and standard deviation of the random variable $|E_c|^2$. Since $|E_c|^2$ is normally distributed, this probability is

$$F\left\{\frac{\mu_c}{\sigma_c}\left[\frac{|\bar{X}-X|^2 \sin^2\theta_o}{\mu_c \sin^2(\alpha-\theta_o)} - 1\right]\right\},$$

where F is the cumulative normal distribution function

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-y^2/2}\, dy .$$

With the previous assumption that the angle α does not depend on the magnitude of E_c, the probability that $\theta \le \theta_o$ is

$$P(\theta \le \theta_o) = P(\alpha \le \theta_o) + \int_{\alpha=\theta_o}^{\alpha=\pi} F\left\{\frac{\mu_c}{\sigma_c}\left[\frac{|\bar{X}-X|^2 \sin^2\theta_o}{\mu_c \sin^2(\alpha-\theta_o)} - 1\right]\right\} p(\alpha)\, d\alpha,$$

where $p(\alpha)$ is the probability density function for the angle α. For values of θ_o such that $P(\alpha \le \theta_o)$ is negligible,

$$P(\theta \le \theta_o) \approx F\left\{\frac{\mu_c}{\sigma_c}\left[\frac{|\bar{X}-X|^2}{\mu_c} \tan^2\theta_o - 1\right]\right\}$$

is a good approximation to the desired probability since $p(\alpha)$ is strongly peaked at $\alpha = \pi/2$ when the dimension of the constraint space is high.

The mean μ_c and standard deviation σ_c of $|E_c|^2$ are fixed for each of the three modes of processing under examination. Given any squared magnitude $|\bar{X}-X|^2$ of the vector $(\bar{X}-X)$, a cumulative distribution function $P(\theta \le \theta_o)$ for the angle θ between the vectors $(\bar{X}-X)$ and $(\bar{X}-X)+E_c$ can be computed. When such distribution functions are evaluated over a range of values of $|\bar{X}-X|^2$,

the result is a probability which is a function of the variables θ_o and $|\bar{X}-X|^2$. Figure III-13 is a perspective drawing of this probability over a rectangular grid of the values θ_o and $|\bar{X}-X|^2$. The probabilities are plotted for the case of unscaled data, data scaled by 16, and floating-point data. In each of the three cases, two drawings are plotted from slightly different viewpoints. The lower plateau in each drawing corresponds to a probability of zero, the upper plateau to a probability of one.

By specifying a probability distribution for the squared magnitudes $|\bar{X}-X|^2$ of the vector $(\bar{X}-X)$, the probability distribution for the angle θ can be found:

$$P(\theta \leq \theta_o) = \int_0^1 P(\theta \leq \theta_o \mid |\bar{X}-X|^2)\, dP\, |\bar{X}-X|^2 \quad .$$

The values $|\bar{X}-X|^2$ computed during a four-hour noise sample from day 238 of 1970 (see Figure III-10) have been assumed to be equally likely. Figure III-14 shows the resulting probability density per degree for the angle θ between the vector $(\bar{X}-X)$ and the projection of the vector $(\bar{X}-X) + E_c + E_b$ onto the constraint space. In the graph corresponding to each of the three modes of processing, there are minor bumps due to local concentrations or local gaps in the values $|\bar{X}-X|^2$.

If the squared magnitudes $|\bar{X}-X|^2$ and $|E_c|^2$ are replaced by their mean values and the angle α is assumed to be 90^o, a useful estimate of the mean $\bar{\theta}$ of the angle θ is obtained:

$$\bar{\theta} \approx \tan^{-1} \sqrt{\frac{\mu_c}{E(|\bar{X}-X|^2)}} \quad ,$$

where $E(|\bar{X}-X|^2)$ is the mean of the squared magnitude $|\bar{X}-X|^2$. For the noise sample from day 238, the mean of $|\bar{X}-X|^2$ was 9693 counts^2

(a) Unscaled Data

(b) Data Scaled by 16

(c) Floating-Point Data

FIGURE III-13

CUMULATIVE DISTRIBUTION FUNCTION $P(\theta \leq \theta_o)$ OVER A RANGE OF THE SQUARED MAGNITUDE $|\overline{X}-X|^2$

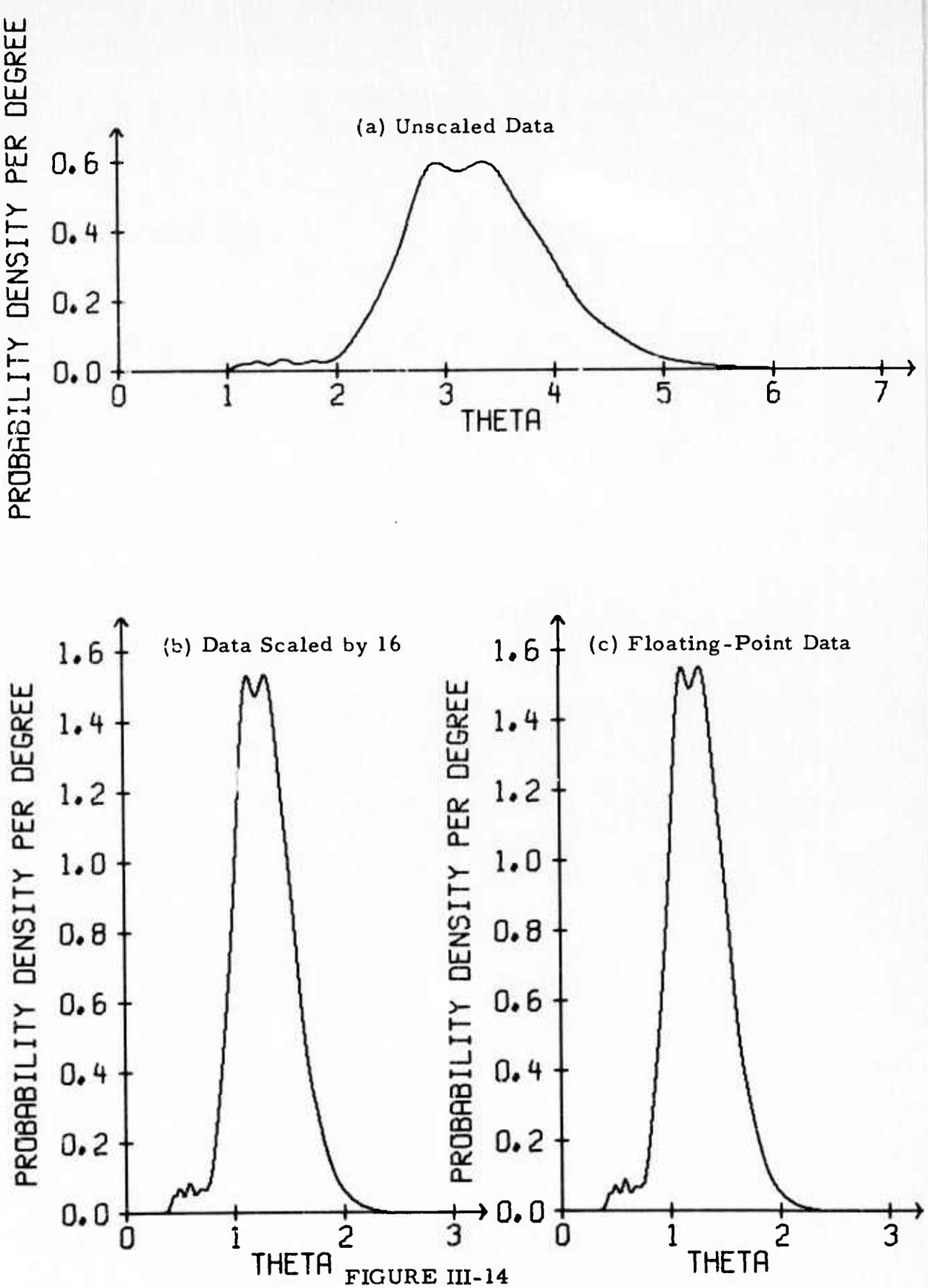


FIGURE III-14

PROBABILITY DENSITY PER DEGREE FOR THE ANGLE BETWEEN THE VECTOR $(\bar{X}-X)$ AND THE PROJECTION OF THE VECTOR $(\bar{X}-X) + E_c + E_b$ ONTO THE CONSTRAINT SPACE

for unscaled data, 2,481,422 counts^2 for data scaled by a factor of 16. Thus the estimated mean of θ is

$$\tan^{-1}\sqrt{\frac{28.071}{9693}} = 3^{o}5',$$

$$\tan^{-1}\sqrt{\frac{1081.952}{2481422}} = 1^{o}12',$$

and

$$\tan^{-1}\sqrt{\frac{4.133}{9693}} = 1^{o}11',$$

respectively, for unscaled data, data scaled by 16, and floating-point data. These estimates are close to the probability density peaks of Figure III-14.

6. Probability Distribution for the Angle ϕ Between the Vectors $(\bar{X}\text{-}X) + E_c + E_b$ and $(\bar{X}\text{-}X)$

As before, it is possible to construct a two-dimensional plane that passes through the origin, any one of the vectors $(\bar{X}\text{-}X)$, and any one of the vectors $(\bar{X}\text{-}X) + E_c + E_b$. Such a plane is shown in Figure III-15. In the figure, ϕ is the angle between the vectors $(\bar{X}\text{-}X) + E_c + E_b$ and $(\bar{X}\text{-}X)$ and E_b is the beamsteer roundoff error vector. $(\bar{X}\text{-}X)$, E_c, and α are as defined previously. The line running from the origin to $(\bar{X}\text{-}X)$ has been extended to the point where $(\bar{X}\text{-}X) + E_c + E_b$ projects onto it in order to form a right triangle. Since E_b is perpendicular to any vector within the constraint space, the leg of the right triangle adjacent to the angle ϕ is of length $|\bar{X}\text{-}X| + |E_c| \cos\alpha$. The side opposite the angle ϕ consists of the sum of the vector E_b and a vector of length $E_c \sin\alpha$ (within the constraint space and perpendicular to $\bar{X}\text{-}X$). Since E_b is perpendicular to the constraint space, the length of the side opposite ϕ is $\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}$. Thus the cotangent of ϕ is

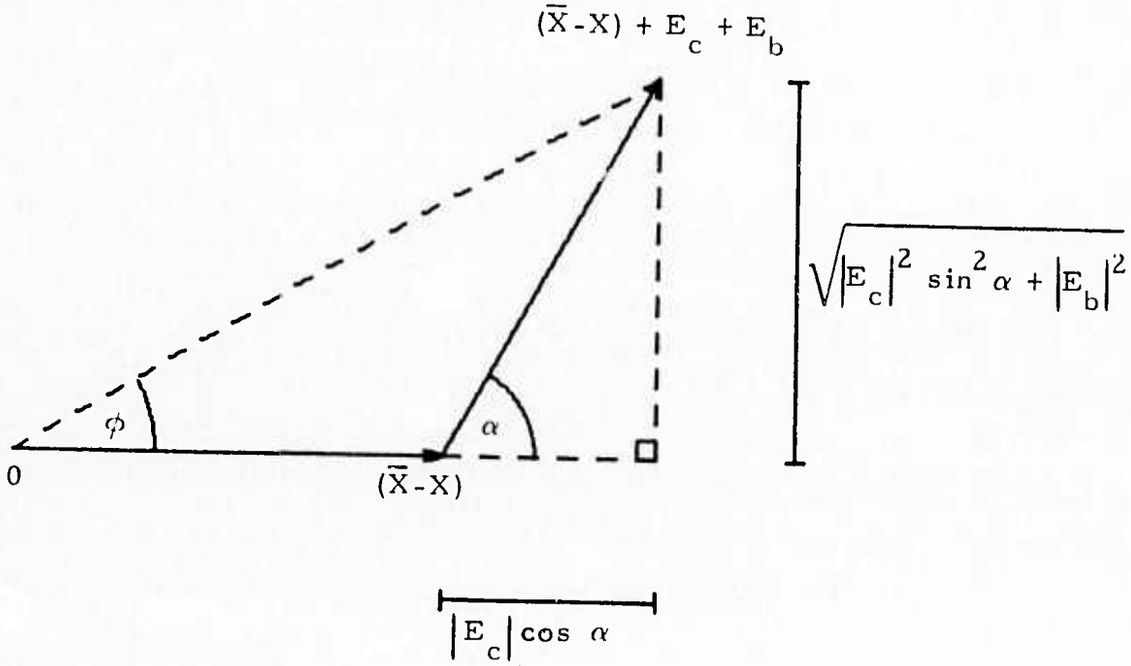


FIGURE III-15

TWO-DIMENSIONAL PLANE PASSING THROUGH THE ORIGIN, $(\bar{X}-X) + E_c + E_b$, AND $(\bar{X}-X)$

$$\cot \phi = \frac{|\bar{X}-X| + |E_c| \cos \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}} .$$

Since the cotangent decreases monotonically in $(0, \pi)$, the angle ϕ is less than or equal to any specific angle ϕ_o within the interval $(0, \pi)$ if and only if $\cot \phi_o \leq \cot \phi$. Hence $\phi \leq \phi_o$ if and only if

$$\cot \phi_o \leq \frac{|\bar{X}-X| + |E_c| \cos \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}} .$$

If $\phi = \phi_o$, $\cot \phi = \cot \phi_o$ implies that

$$|E_c|^2 \sin^2 \alpha + |E_b|^2 = (|\bar{X}-X| + |E_c| \cos \alpha)^2 \tan^2 \phi_o .$$

This equation is consistent with the conic section corresponding to the intersection of a cone of angle ϕ_o and a cutting plane perpendicular to the constraint space. The cone of angle ϕ_o has its vertex at the origin and is centered about an axis through the origin and $(\bar{X}-X)$. The cutting plane is determined by the three points $(\bar{X}-X)$, $(\bar{X}-X) + E_c$, and $(\bar{X}-X) + E_c + E_b$. For a given value of $|\bar{X}-X|^2$ and a given angle α, the region where $\phi \leq \phi_o$ can be expressed as an inequality involving the two variables $|E_c|$ and $|E_b|$. In fact, if $(\bar{X}-X)$ is selected as the origin of a coordinate system within the cutting plane, the direction of the vectors E_c and E_b can be used to specify the directions of the abscissa and ordinate, respectively, in a two-dimensional Cartesian system with perpendicular axes. For example, if $\phi_o < \pi/2 = \alpha$, the region where $\phi \leq \phi_o$ can be expressed in terms of the inequality

$$|E_b|^2 + |E_c|^2 \leq |\bar{X}-X|^2 \tan^2 \phi_o ,$$

which corresponds to the interior of a circle of radius $|\bar{X}-X| \tan \phi_o$ with its center at the origin $(\bar{X}-X)$ of the cutting-plane coordinate system. Since the axis directions are determined by E_c and E_b, the coordinates are $(|E_c|, |E_b|)$ for any vector pair (E_c, E_b). Thus all points lie within the first quadrant of the new coordinate system for the cutting plane.

It is easier to define the region where $\phi \leq \phi_o$ with inequalities which express $|E_b|$ in terms of $|E_c|$. Since the results obtained in this way will help to specify $|E_c|$ in terms of $|E_b|$, they will be derived first.

If $\phi_o = 0$, $\phi \leq \phi_o$ implies that $\phi = \phi_o = 0$ and

$$\tan \phi = 0 = \frac{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}}{|\bar{X}-X| + |E_c| \cos \alpha}$$

Thus $|E_b|^2$ and $|E_c|^2 \sin^2 \alpha$ must be zero. If $0 < \alpha < \pi$, only the single point $(|E_c|, |E_b|) = (0, 0)$ satisfies the requirement that $\phi \leq \phi_o$. If $\alpha = 0$, $|E_b|$ must be zero and all values of $|E_c|$ must lie on the semi-infinite line $0 \leq |E_c| \leq \infty$. If $\alpha = \pi$, $|E_b|$ must again be zero and all values of $|E_c|$ must lie on the line segment $0 \leq E_c \leq |\bar{X}-X|$. In this case, the angle ϕ is undefined when $|E_c| = |\bar{X}-X|$.

If $0 < \phi_o < \pi/2$ and $\alpha \leq \phi_o$,

$$\tan^2 \phi_o \left[|\bar{X}-X| + 2|\bar{X}-X||E_c| \cos \alpha + |E_c|^2 \sin^2 \alpha (\cot^2 \alpha - \cot^2 \phi_o)\right]$$

$$= (|\bar{X}-X| + |E_c| \cos \alpha)^2 \tan^2 \phi_o - |E_c|^2 \sin^2 \alpha \geq 0 .$$

The required inequality

$$\cot \phi_o = \frac{|\bar{X}-X| + |E_c| \cos \alpha}{(|\bar{X}-X| + |E_c| \cos \alpha) \tan \phi_o} \leq \frac{|\bar{X}-X| + |E_c| \cos \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}}$$

is satisfied if and only if

$$0 \le |E_b|^2 \le (|\bar{X}-X| + |E_c| \cos\alpha)^2 \tan^2\phi_o - |E_c|^2 \sin^2\alpha .$$

If $0 < \phi_o < \pi/2$ and $\alpha > \phi_o$, φ is always greater than ϕ_o if $|E_c| > |\bar{X}-X| \sin\phi_o / \sin(\alpha-\phi_o)$. The following inequalities are valid if $\sin\alpha > 0$:

$$|E_c| \sin\alpha > \frac{|\bar{X}-X| \sin\alpha \sin\phi_o}{\sin(\alpha-\phi_o)} = \frac{|\bar{X}-X|}{\cot\phi_o - \cot\alpha} ;$$

$$\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}\ \cot\phi_o \ge |E_c| \sin\alpha \cot\phi_o > |\bar{X}-X| + |E_c| \cos\alpha;$$

$$\cot\phi_o > \frac{|\bar{X}-X| + |E_c| \cos\alpha}{\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}} .$$

The last inequality is still valid if $|E_b| > 0$. If $|E_b| = 0$ and $\alpha = \pi$, $|E_c| > |\bar{X}-X|$ implies that $|\bar{X}-X| + |E_c| \cos\pi < 0$ and hence $\phi = \pi$. Now if $|E_c| \le |\bar{X}-X| \sin\phi_o / \sin(\alpha-\phi_o)$,

$$0 \le |E_c| \sin\alpha \le (|\bar{X}-X| + |E_c| \cos\alpha) \tan\phi_o$$

and

$$(|\bar{X}-X| + |E_c| \cos\alpha)^2 \tan^2\phi_o - |E_c|^2 \sin^2\alpha \ge 0 .$$

Accordingly, when $|E_c| \le |\bar{X}-X| \sin\phi_o / \sin(\alpha-\phi_o)$,

$$\cot\phi_o = \frac{|\bar{X}-X| + |E_c| \cos\alpha}{(|\bar{X}-X| + |E_c| \cos\alpha)\tan\phi_o} \le \frac{|\bar{X}-X| + |E_c| \cos\alpha}{\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}}$$

if and only if

$$0 \le |E_b|^2 \le (|\bar{X}-X| + |E_c| \cos \alpha)^2 \tan^2 \phi_o - |E_c|^2 \sin^2 \alpha .$$

When $0 < \phi_o < \pi/2$ and $\alpha > \phi_o$, therefore, $\phi \le \phi_o$ if and only if both of the conditions

$$|E_c| \le \frac{|\bar{X}-X| \sin \phi_o}{\sin(\alpha - \phi_o)}$$

$$0 \le |E_b|^2 \le (|\bar{X}-X| + |E_c| \cos \alpha)^2 \tan^2 \phi_o - |E_c|^2 \sin^2 \alpha$$

are met.

If $\phi_o = \pi/2$,

$$0 = \cot \phi_o \le \frac{|\bar{X}-X| + |E_c| \cos \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}}$$

if and only if

$$|\bar{X}-X| + |E_c| \cos \alpha \ge 0 .$$

If $\alpha \le \pi/2$, the condition is always satisfied. If $\alpha > \pi/2$, $\phi \le \phi_o$ if and only if

$$|E_c| \le \frac{|\bar{X}-X|}{-\cos \alpha} = \frac{|\bar{X}-X| \sin \phi_o}{\sin(\alpha - \phi_o)} .$$

If $\pi/2 < \phi_o \le \pi$ and $\alpha \le \phi_o$, ϕ is never greater than ϕ_o. Either $\alpha \le \pi/2$ and $\cot \alpha \ge 0$ so that

$$\cot \phi_o - \left(\frac{|E_c| \sin \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}} \right) \cot \alpha < 0 \le \frac{|\bar{X}-X|}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}}$$

and

$$\cot \phi_o < \frac{|\bar{X}-X| + |E_c| \cos \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}}$$

or $\alpha > \pi/2$ and $\cot \alpha < 0$ so that

$$\cot \phi_o + \left(\frac{|E_c| \sin \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}} \right) (-\cot \alpha)$$

$$\leq \cot \phi_o - \cot \alpha \leq 0 \leq \frac{|\bar{X}-X|}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}}$$

and

$$\cot \phi_o \leq \frac{|\bar{X}-X| + |E_c| \cos \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}} .$$

Finally, if $\pi/2 < \phi_o \leq \pi$ and $\alpha > \phi_o$, ϕ is never greater than ϕ_o if $|E_c| \leq |\bar{X}-X| \sin \phi_o / \sin(\alpha - \phi_o)$. The following conditions are satisfied:

$$|E_c| \sin \alpha \leq \frac{|\bar{X}-X| \sin \alpha \sin \phi_o}{\sin(\alpha - \phi_o)} = \frac{|\bar{X}-X|}{\cot \phi_o - \cot \alpha} ;$$

$$\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2} \cot \phi_o \leq |E_c| \sin \alpha \cot \phi_o \leq |\bar{X}-X| + |E_c| \cos \alpha;$$

$$\cot \phi_o \leq \frac{|\bar{X}-X| + |E_c| \cos \alpha}{\sqrt{|E_c|^2 \sin^2 \alpha + |E_b|^2}} .$$

On the other hand, if $|E_c| > |\bar{X}-X| \sin\phi_o / \sin(\alpha-\phi_o)$, then $(|\bar{X}-X| + |E_c|\cos\alpha)^2 \tan^2\phi_o - |E_c|^2 \sin^2\alpha \geq 0$ and

$$-\cot\phi_o = \frac{-(|\bar{X}-X| + |E_c|\cos\alpha)}{(|\bar{X}-X| + |E_c|\cos\alpha)\tan\phi_o} \geq \frac{-(|\bar{X}-X| + |E_c|\cos\alpha)}{\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}}$$

if and only if

$$|E_b|^2 \geq (|\bar{X}-X| + |E_c|\cos\alpha)^2 \tan^2\phi_o - |E_c|^2 \sin^2\alpha .$$

Since $|E_b|^2$ has a discrete probability distribution and $|E_c|^2$ is normally distributed, it is easier to compute the probability that $\phi \leq \phi_o$ when $|E_c|$ is given in terms of $|E_b|$. For each value of $|E_b|$, there exists a range of $|E_c|$ values such that $\phi \leq \phi_o$. The probability of this range of $|E_c|$ values is the difference between two cumulative normal distribution functions.

To specify the boundary of the region where $\phi \leq \phi_o$, the quadratic equation

$$(|E_c|^2 \sin^2\alpha + |E_b|^2)\cot^2\phi_o = (|\bar{X}-X| + |E_c|\cos\alpha)^2$$

is used to solve $|E_c|$ in terms of $|E_b|$. Ordinarily only one solution for $|E_c|$ is the correct definition of the boundary. Since the conditions for $\phi = 0$ have negligible probability ($|E_b|^2 = 0$) and since $\phi \leq \phi_o$ whenever $\pi/2 \leq \phi_o$ and $\alpha \leq \phi_o$, only the cases $0 < \phi_o < \pi/2$ and $\pi/2 \leq \phi_o < \alpha$ will be discussed.

When $0 < \phi_o < \pi/2$ and $\alpha \leq \phi_o$, the inequality $|E_b| \leq |\bar{X}-X| \tan\phi_o$ guarantees that

$$(|\bar{X}-X| + |E_c|\cos\alpha)\tan\phi_o \geq \sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}$$

and hence that $\varphi \leq \phi_o$:

$$\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2} \leq \sqrt{|E_c|^2 \cos^2\alpha \tan^2\alpha + |\bar{X}-X|^2 \tan^2\phi_o}$$

$$\leq \tan\phi_o \sqrt{|E_c|^2 \cos^2\alpha + |\bar{X}-X|^2} \leq \tan\phi_o \quad (|\bar{X}-X| + |E_c| \cos\alpha).$$

If $|E_b| > |\bar{X}-X| \tan\phi_o$ when $0 < \phi_o < \pi/2$ and $\alpha \leq \phi_o$, $\varphi \leq \phi_o$ if and only if $|E_c|$ satisfies the linear inequality

$$|E_c| \tan\phi_o \geq |E_b| - |\bar{X}-X| \tan\phi_o$$

when $\alpha = 0$, the hyperbolic inequality

$$|E_c| \sin\alpha \geq \frac{\cot\phi_o \sqrt{|\bar{X}-X|^2 + |E_b|^2 (\cot^2\alpha - \cot^2\phi_o)} - |\bar{X}-X| \cot\alpha}{\cot^2\alpha - \cot^2\phi_o}$$

when $0 < \alpha < \phi_o$, or the parabolic inequality

$$|E_c| \cos\alpha \geq \frac{|E_b|^2 \cot^2\phi_o - |\bar{X}-X|^2}{2|\bar{X}-X|}$$

when $\alpha = \phi_o$.

If $0 < \phi_o < \pi/2$ and $\alpha > \phi_o$, $|E_c|$ must be no larger than $|\bar{X}-X| \sin\phi_o / \sin(\alpha - \phi_o)$ in order that $\varphi \leq \phi_o$ (as was shown earlier). The most difficult case occurs when $\phi_o < \alpha < \pi/2$. Whenever $|E_b|^2 > |\bar{X}-X|^2/(\cot^2\phi_o - \cot^2\alpha)$, ϕ is always greater than ϕ_o :

$$(|E_c|^2 \sin^2\alpha + |E_b|^2) \cot^2\varphi_o$$

$$> \left(|E_c|^2 \sin^2\alpha + \frac{|\bar{X}-X|^2}{\cot^2\phi_o - \cot^2\alpha} \right) \cot^2\phi_o$$

$$= |E_c|^2 \cos^2\alpha + |E_c|^2 \sin^2\alpha \,(\cot^2\phi_o - \cot^2\alpha) + \frac{|\bar{X}-X|^2 \cot^2\alpha}{\cot^2\phi_o - \cot^2\alpha} + |\bar{X}-X|^2$$

$$\geq |E_c|^2 \cos^2\alpha + 2|E_c|\cos\alpha\,|\bar{X}-X| + |\bar{X}-X|^2 = (|\bar{X}-X| + |E_c|\cos\alpha)^2$$

since

$$0 \leq (\cot^2\phi_o - \cot^2\alpha) \left[|E_c| \sin\alpha - \frac{|\bar{X}-X|\cot\alpha}{\cot^2\phi_o - \cot^2\alpha} \right]^2$$

$$= |E_c|^2 \sin^2\alpha\,(\cot^2\phi_o - \cot^2\alpha) - 2|E_c|\cos\alpha\,|\bar{X}-X| + \frac{|\bar{X}-X|^2 \cot^2\alpha}{\cot^2\phi_o - \cot^2\alpha}.$$

Thus $\sqrt{|E_c|^2 \sin^2 + |E_b|^2}\,\cot\phi > |\bar{X}-X| + |E_c|\cos\alpha$ and $\phi > \phi_o$. For the range $0 \leq |E_b|^2 \leq |\bar{X}-X|^2 \tan^2\phi_o$, $\phi \leq \phi_0$ if and only if

$$0 \leq |E_c| \sin\alpha \leq \frac{|\bar{X}-X|\cot\alpha + \cot\phi_o \sqrt{|\bar{X}-X|^2 - |E_b|^2(\cot^2\phi_o - \cot^2\alpha)}}{\cot^2\phi_o - \cot^2\alpha}$$

when $\phi_o < \alpha < \pi/2$. For the range $|\bar{X}-X|^2 \tan^2\phi_o \leq |E_b|^2 \leq \frac{|\bar{X}-X|^2}{\cot^2\phi_o - \cot^2\alpha}$, both of the inequalities

$$\frac{|\bar{X}-X|\cot\alpha - \cot\phi_o \sqrt{|\bar{X}-X|^2 - |E_b|^2(\cot^2\phi_o - \cot^2\alpha)}}{\cot^2\phi_o - \cot^2\alpha} \leq |E_c|\sin\alpha$$

and

$$|E_c| \sin\alpha \leq \frac{|\bar{X}-X|\cot\alpha + \cot\varphi_o \sqrt{|\bar{X}-X|^2 - |E_b|^2 (\cot^2\varphi_o - \cot^2\alpha)}}{\cot^2\varphi_o - \cot^2\alpha}$$

must be satisfied if $\phi \leq \phi_o$. The region $\varphi \leq \varphi_o$ forms an ellipse in the cutting plane when $0 < \phi_o < \alpha < \pi/2$.

If $0 < \varphi_o < \pi/2 \leq \alpha$, ϕ is always greater than ϕ_o whenever $|E_b| > |\bar{X}-X|\tan\phi_o$:

$$\cot\varphi_o \geq \frac{|\bar{X}-X|}{\sqrt{|E_c|^2 \sin^2\alpha + |\bar{X}-X|^2 \tan^2\phi_o}}$$

$$> \frac{|\bar{X}-X|}{\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}} \geq \frac{|\bar{X}-X| + |E_c|\cos\alpha}{\sqrt{|E_c|^2 \sin^2\alpha + |E_b|^2}} .$$

If $|E_b| \leq |\bar{X}-X|\tan\varphi_o$ when $0 < \varphi_o < \pi/2 \leq \alpha$, $\varphi \leq \varphi_o$ if and only if $|E_c|$ satisfies the circular inequality

$$|E_c|^2 \leq |\bar{X}-X|^2 \tan^2\phi_o - |E_b|^2$$

when $\alpha = \pi/2$, the elliptic inequality

$$|E_c| \sin\alpha \leq \frac{\cot\varphi_o \sqrt{|\bar{X}-X|^2 - |E_b|^2 (\cot^2\phi_o - \cot^2\alpha)} + |\bar{X}-X|\cot\alpha}{\cot^2\phi_o - \cot^2\alpha}$$

when $\frac{\pi}{2} < \alpha < \pi - \phi_o$, the parabolic inequality

$$|E_c|(-\cos\alpha) \leq \frac{|\bar{X}-X|^2 - |E_b|^2 \cot^2 \phi_o}{2|\bar{X}-X|}$$

when $\alpha = \pi - \phi_o$, the hyperbolic inequality

$$|E_c|\sin\alpha \leq \frac{-|\bar{X}-X| \cot\alpha - \cot\phi_o \sqrt{|\bar{X}-X|^2 + |E_b|^2(\cot^2\alpha - \cot^2\phi_o)}}{\cot^2\alpha - \cot^2\phi_o}$$

when $\pi - \phi_o < \alpha < \pi$, or the linear inequality

$$|E_c| \tan\phi_o \leq |\bar{X}-X| \tan\phi_o - |E_b|$$

when $\alpha = \pi$.

It was shown earlier that $\phi \leq \phi_o$ if and only if

$$|E_c|(-\cos\alpha) \leq |\bar{X}-X|$$

when $\pi/2 = \phi_o < \alpha$.

If $\pi/2 < \phi_o < \alpha$, $\phi \leq \phi_o$ if and only if $|E_c|$ satisfies the hyperbolic inequality

$$|E_c| \sin\alpha \leq -\left[\frac{|\bar{X}-X| \cot\alpha + \cot\phi_o \sqrt{|\bar{X}-X|^2 + |E_b|^2(\cot^2\alpha - \cot^2\phi_o)}}{\cot^2\alpha - \cot^2\phi}\right]$$

when $\alpha < \pi$, or the linear inequality

$$|E_c| \leq |\bar{X}-X| - |E_b| \cot\phi_o$$

when $\alpha = \pi$. These last two inequalities are automatically satisfied when $|E_c| \leq |\bar{X}-X| \sin\phi_o / \sin(\alpha - \phi_o)$.

Figure III-16 shows the region in the cutting plane where $\phi \leq \phi_o$ for a variety of angles α and ϕ_o. The normalized variables $C = |E_c| / |\bar{X}-X|$ and $B = |E_b| / |\bar{X}-X|$ replace the variables $|E_c|$ and $|E_b|$, respectively.

The most easily obtained information about the probability distribution of ϕ is an estimate of its mean $\bar{\phi}$. If the squared magnitudes $|\bar{X}-X|^2$, $|E_c|^2$, and $|E_b|^2$ are replaced by their mean values and the angle α is assumed to be 90^o,

$$\bar{\phi} \approx \tan^{-1} \sqrt{\frac{\mu_c + \mu_b}{E(|\bar{X}-X|^2)}},$$

where $E(|\bar{X}-X|^2)$ is the mean of the squared magnitude $|\bar{X}-X|^2$, μ_c is the mean of $|E_c|^2$, and μ_b is the mean of $|E_b|^2$. For the noise sample from day 238 of 1970 the estimated mean of ϕ is

$$\tan^{-1} \sqrt{\frac{28.071 + 16.361}{9693}} = 3^o\ 52',$$

$$\tan^{-1} \sqrt{\frac{1081.952 + 16.361}{2481422}} = 1^o\ 12',$$

$$\tan^{-1} \sqrt{\frac{4.133 + 0.0}{9693}} = 1^o\ 11',$$

respectively, for unscaled data, data scaled by 16, and floating-point data. The estimate of the mean for the angle of error in $(\bar{X}-X)$ is 47' greater for unscaled data, less than 1' greater for data scaled by 16, and, of course, exactly the same for floating-point data when the effect of the beamsteer roundoff error is considered.

The associated equation

$$E(|\bar{X}-X|^2) \approx \frac{\mu_c + \mu_b}{\tan^2 \bar{\phi}}$$

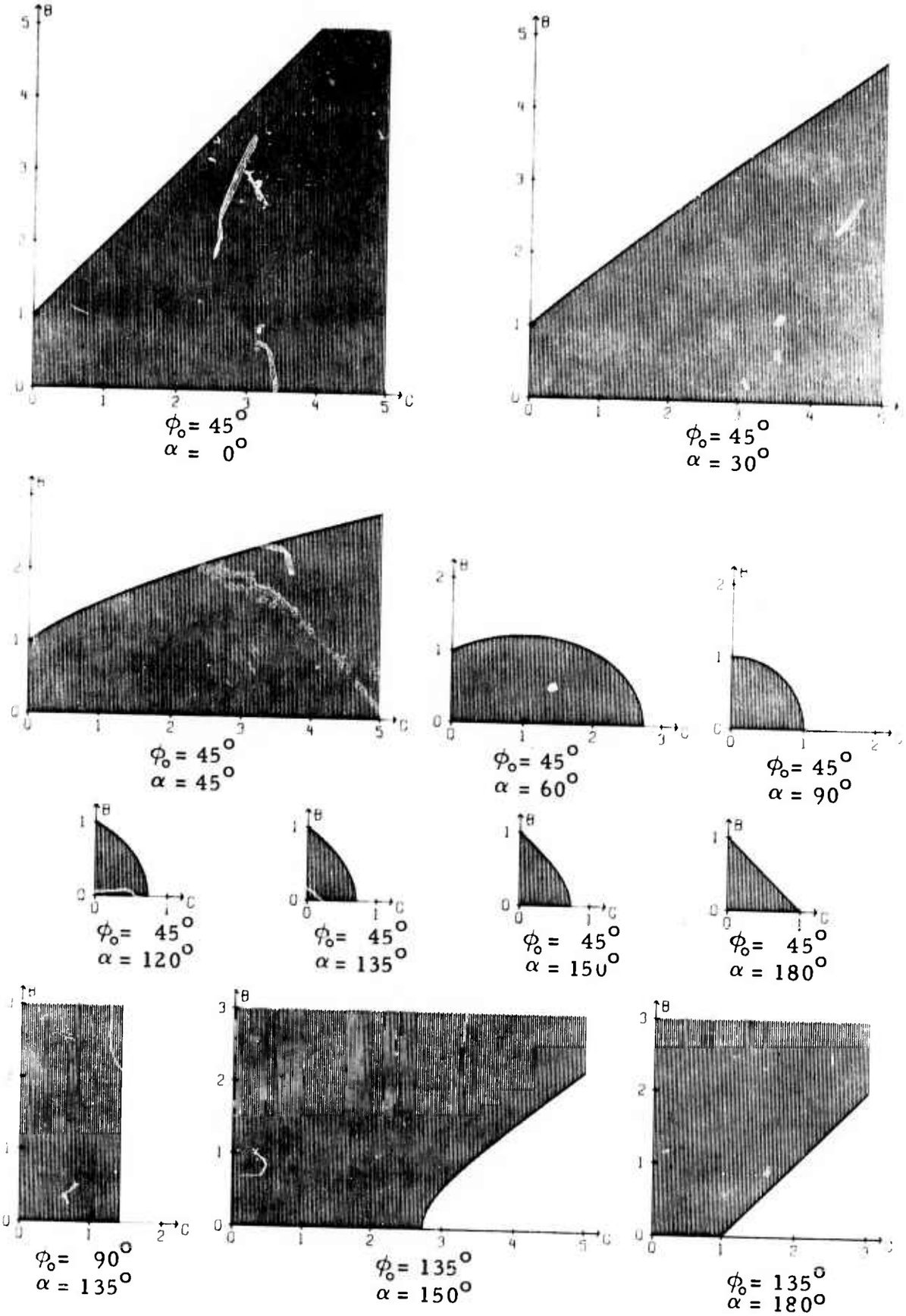


FIGURE III-16

REGION IN THE CUTTING PLANE WHERE $\phi \leq \phi_o$ IN TERMS OF THE NORMALIZED CUTTING PLANE COORDINATES $C = |E_c| / |\bar{X}-X|$ AND $B = |E_b| / |\bar{X}-X|$

can be used to select analog-to-digital conversion factors and computer data scale factors which achieve a desired angle of error $\bar{\phi}$. If the A/D system is fixed, for example, the minimum achievable angular error can be determined by considering only the digitization errors. This minimum achievable error can be approached arbitrarily closely through the choice of a suitably high data scale factor. As the data scale factor increases, so does the term $E(|\bar{X}-X|^2)$ when expressed in squared computer counts. In turn, the computer-count value of $E(|\bar{X}-X|)$ is almost sufficient to express the computer-count RMS level for the data points $x_i(t-j)$ and hence the number of bits required to express it. The dynamic range of the data is a separate problem, however. The number of guard bits needed to reduce the frequency of data clipping to acceptable levels is a function of the fluctuations above the data RMS level. If both the A/D system and computer data scaling are subject to control, a wider range of A/D conversion factors and computational scale factors is possible. For a given desired angle $\bar{\phi}$, in this event, there will be a curve relating the A/D conversion factor to the computer data scale factor.

The more useful approximation

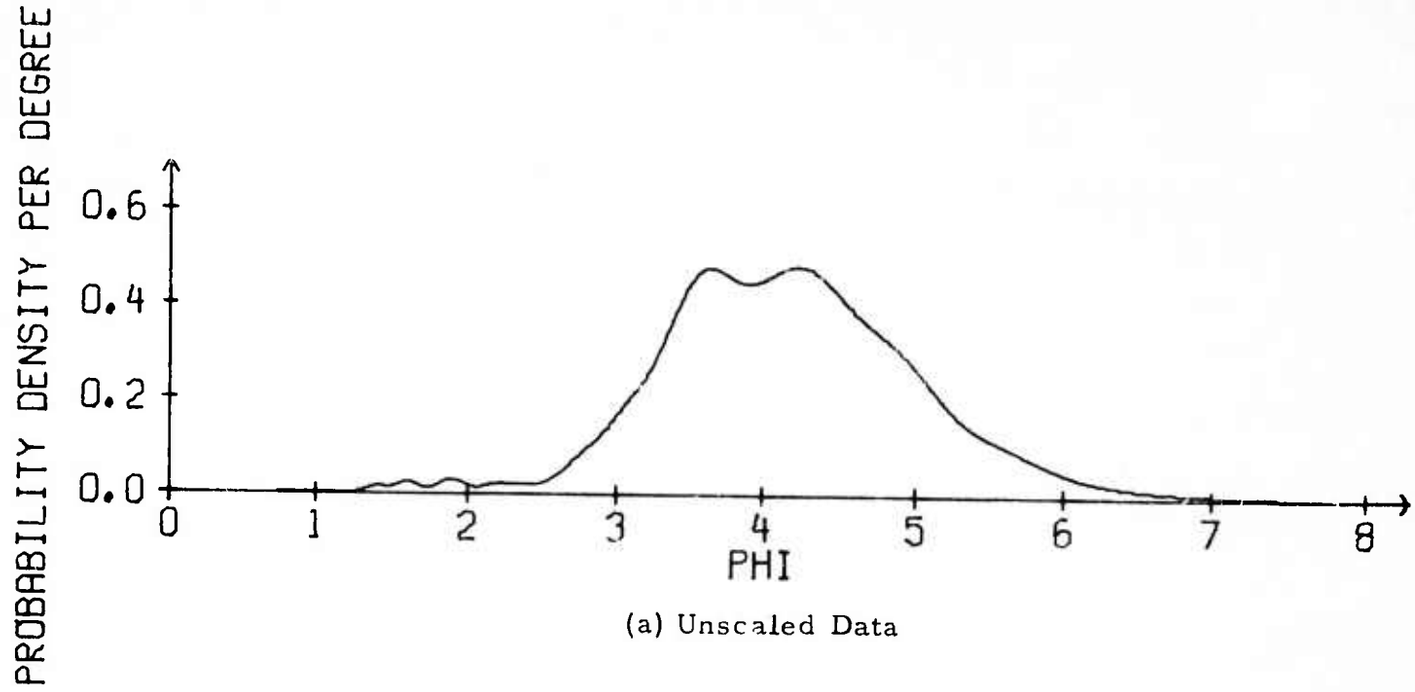
$$P(\phi \leq \phi_o) \approx P\left(|E_c|^2 + |E_b|^2 \leq |\bar{X}-X|^2 \tan^2 \phi_o\right)$$

is based on the assumption that the angle α is always $\pi/2$. Since the probability density for α is strongly and symmetrically concentrated about a value of $\pi/2$ and since the probability that $\phi \leq \phi_o$ is quasi-linear in the neighborhood of $\alpha = \pi/2$, the approximation is remarkably accurate. Along the curve $|\bar{X}-X|^2 \tan^2 \phi_o = \mu_c$ (where the estimate is 1/2 when $|E_b|^2 = 0$), the actual values ranged from 0.50 at $\alpha = \pi/2$ to less than 0.55 at the highest $|\bar{X}-X|^2$ value. In most cases the approximation underestimates $P(\phi \leq \phi_o)$ and hence tends to overestimate the angle of error. By specifying a probability distribution for the squared magnitudes $|\bar{X}-X|^2$, the probability distribution for the angle ϕ can be approximated:

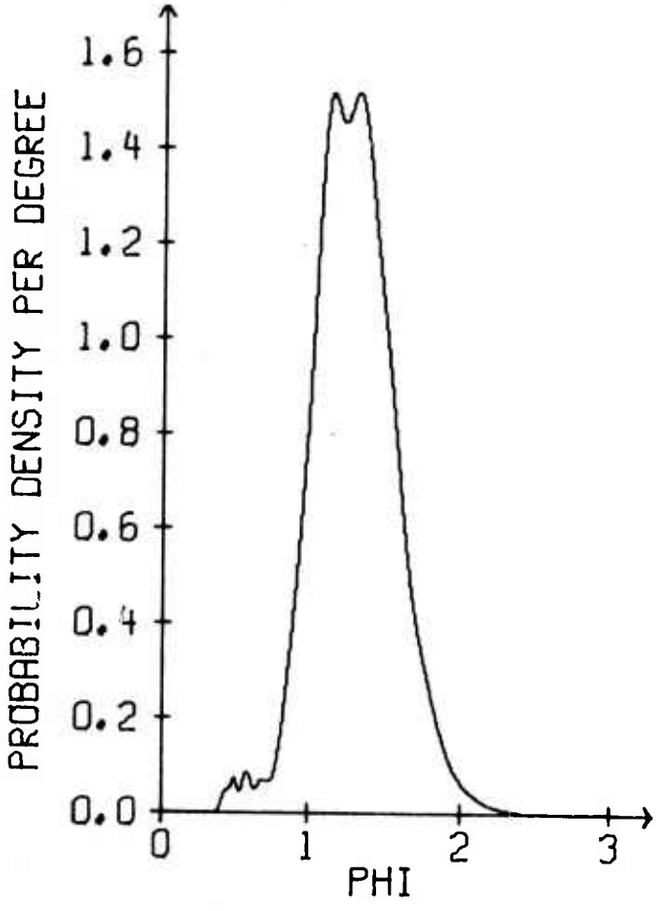
$$P(\varphi \leq \varphi_o) \approx \int_0^1 P(|E_c|^2 + |E_b|^2 \leq |\bar{X}-X|^2 \tan^2 \phi_o)\, dP|\bar{X}-X|^2 .$$

Once again, as in the case of the probability density functions graphed in Figure III-14, the values $|\bar{X}-X|^2$ from a four-hour noise sample from day 238 of 1970 have been assumed equally likely. Figure III-17 is a plot of the corresponding probability density per degree for the angle of error φ. Some idea of the accuracy of the approximation can be gleaned from a comparison of Figures III-14 (c) and III-17 (c). In the case of floating-point data, $|E_b|^2 = 0$ and the two probability distributions are identical: the approximation has shifted the twin peaks at 1^o 7.5' and 1^o 17.25' to 1^o 7.8' and 1^o 17.55' (only 0.3' error). In the case of data scaled by 16, the twin peaks in Figure III-14 (b) are 0.9' higher than for the floating-point data. After allowing for error perpendicular to the constraint space, they are now 1.2' to 1.5' higher in Figure III-17 (b) than the peaks for floating-point data. With unscaled data, however, noticeable increases are visible in Figure III-17 (a) as compared with Figure III-14 (a). The first peak has moved 44.25' from 2^o 55.5' to 3^o 39.75', the second 53.25' from 3^o 21.375' to 4^o 14.625'. These changes (due to beamsteer round-off error) are in rough agreement with the 47' increase predicted by the approximations for $\bar{\theta}$ and $\bar{\phi}$.

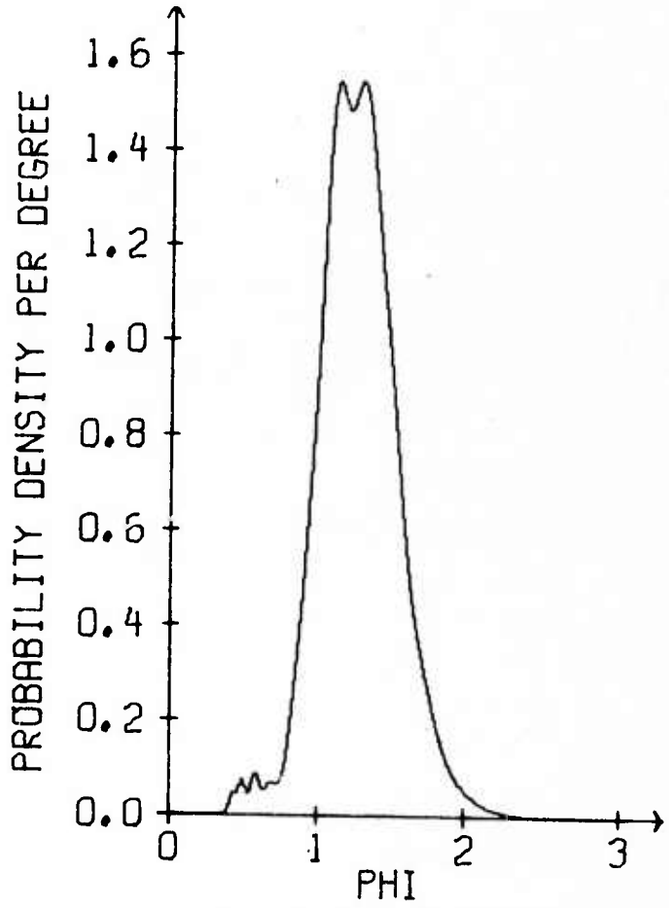
If the complete calculation is to be performed, probabilities must be computed over a range of $|\bar{X}-X|^2$ and ϕ_o values. To obtain one of these probabilities, a numerical integration over the interval $0 \leq \alpha \leq \pi$ is necessary. And, finally, for any particular angle α, a summation of cumulative normal distribution functions (each term corresponding to one discrete value of $|E_b|^2$) must be carried out. The necessary mathematical relationships have been detailed in this part of subsection C. Although it is possible in principle to accomplish the required computations, they have not been done because of the computer time involved. It is necessary to be content with the results displayed in Figure III-17.



(a) Unscaled Data



(b) Data Scaled by 16



(c) Floating-Point Data

FIGURE III-17

PROBABILITY DENSITY PER DEGREE FOR THE ANGLE ϕ BETWEEN THE VECTOR $(\bar{X}-X)$ AND THE VECTOR $(\bar{X}-X) + E_c + E_b$

At this point, it is worth mentioning that the beamsteer roundoff error can be eliminated by multiplying each component $\bar{x}(t-j) - x_i(t-j)$ in the vector $(\bar{X}-X)$ by M, the number of channels:

$$M\left[\bar{x}(t-j) - x_i(t-j)\right]$$

$$= \left[\sum_{k=1}^{M} x_k(t-j)\right] - M\,x_i(t-j) \quad .$$

In calculating the filter change vector $(A^{old} - A^{new})$, a compensatory scaling by $1/M$ can be executed. In the scaled vector $M(\bar{X}-X)$, all the quantities involved are integers. The result is an integer, and no roundoff error is produced in the scaled beamsteer output $M\,\bar{x}(t-j)$. If no overflow occurs, the scaled vector $M(\bar{X}-X)$ satisfies the constraint conditions exactly. The eradication of the beamsteer roundoff entails one multiplication for each point $x_i(t-j)$ used in the adaptive filter update. If the values $M\,x_i(t-j)$ are stored for use in subsequent updates, only M multiplications per sample period are required. Even so, however, the additional computations may preclude implementation of this procedure.

The discussion near the beginning of subsection B mentioned that noise reduction for one data sample was increased from 2 dB to 6 dB (relative to the beamsteer output) simply by scaling the data points by a factor of 16. In view of the relatively small angle of error in the vector $\bar{X}-X$ (even for unscaled data), the principal source of error probably occurs elsewhere.

D. ERROR IN THE FILTER OUTPUT $X^T A = y(t)$

In this subsection, the effects on the adaptive filter output $y(t)$ of error in the data vector X are studied. One effect is that the sign of the adaptive filter output may change so that the filter step

$$A^{new} - A^{old} = \frac{2K_s\, y(t)\,(\bar{X}-X)}{(\bar{X}-X)^T(\bar{X}-X)}$$

is in the wrong direction. As mentioned in subsection A, the effect of reversing the direction of the filter update vector is the precise opposite of what adaptive filtering sets out to accomplish: instead of reducing the mean square filter output $y^2(t)$, the sign reversal causes the filter vector to move in a direction which tends to increase the average squared filter output. A second effect is that the filter vector A does not move if the dot product $(X + E_c)^T A$ lies between $-1/2$ and $1/2$ and is consequently rounded to zero when integer arithmetic is used. (This effect does not take place when floating-point arithmetic is employed.) When the filter update vector cannot move, the adaptive-filtering process is delayed until the next non-zero value $y(t)$ of the filter output. Over a large number of iterations, the result is an effective reduction in the convergence rate through neglect of those times t where the adaptive filter output $y(t)$ is zero. These first two phenomena are discussed in part 1. A completely different effect is that the filter output power is altered. A method to calculate the modified filter output power from the original crosscorrelation matrix (without digitization and computational error) and the probability distributions for the error is derived in part 2.

1. Probability That $y(t)$ Changes Sign or Rounds To Zero

In order to calculate the probability that the filter output $y(t)$ changes sign or rounds to zero after adding the error vector E_c to the data vector X, the probability distribution for the error in $y(t)$

$$\epsilon_y(t) = \sum_{j=-N}^{N} \sum_{i=1}^{M} a_i(j)\ \epsilon_i(t-j)$$

before roundoff must be determined. Since the probability density function

for the roundoff error $\epsilon_i(t-j)$ in any single component of E_c is symmetric about zero, the probability distribution for the error $\epsilon_y(t)$ in y(t) before roundoff is the same as for the random variable

$$\sum_{j=-N}^{N} \sum_{i=1}^{M} |a_i(j)| \, \epsilon_i(t-j) \quad .$$

The probability density function for this random variable is the convolution of the probability density functions for the M(2N+1) random variables $|a_i(j)|\epsilon_i(t-j)$. The variance of $\epsilon_y(t)$ is least when the adaptive filter output is the beamsteer output

$$\bar{x}(t) = \frac{1}{M} \sum_{i=1}^{M} x_i(t)$$

and greatest when each filter weight $a_i(j)$ assumes its maximum absolute value. In the adaptive filtering program used for this report, the filter weights clip when their absolute value is one half, so that the highest variance occurs when

$$\epsilon_y(t) = \frac{1}{2} \sum_{j=-N}^{N} \sum_{i=1}^{M} \epsilon_i(t-j) \quad .$$

Such a situation can happen when, at the zero lag j=0, four weights are 1/2 two -1/2, and, at non-zero lags $j \neq 0$, three weights are 1/2, three -1/2. The probability density function for the most favorable case, i.e., when $y(t)=\bar{x}(t)$, was plotted in Figure III-3. When, in the most unfavorable case conceivable, all 186 filter weights are of magnitude one half, the probability distribution for the random variable

$$\frac{1}{2} \sum_{j=-N}^{N} \sum_{i=1}^{M} \epsilon_i(t-j)$$

is accurately approximated by a normal distribution, so that the probability density function for the error in $y(t)$ before roundoff is

$$p\left[\epsilon_y(t)\right] = \frac{1}{\sigma_\epsilon \sqrt{2\pi}} \quad e^{-(\epsilon^2/2\sigma_\epsilon^2)} ,$$

where σ_ϵ^2 is the sum of the 186 identical variances for each of the random variables $\epsilon_i(t-j)/2$. That is to say,

$$\sigma_\epsilon^2 = 46.5 \text{ Var}\left[\epsilon_i(t-j)\right] .$$

The resulting probability density functions for the three cases are graphed in Figure III-18.

It is now possible to calculate the conditional probability that the sign of $y(t)$ is inverted given the absolute value $|y(t)|$: it is

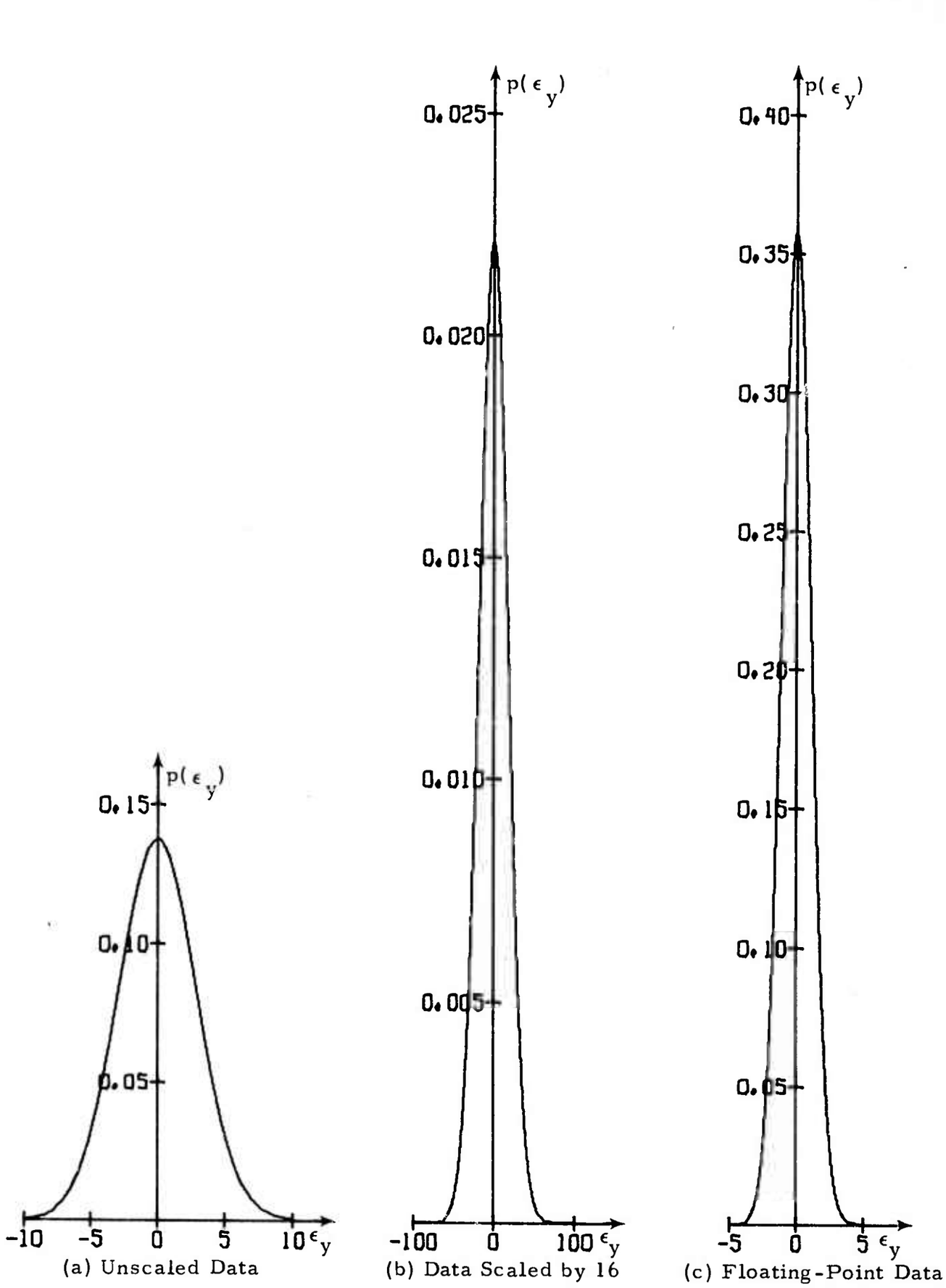
$$\int_{|y(t)| + 1/2}^{\infty} p\left[\epsilon_y(t)\right] d\epsilon_y(t)$$

for integer arithmetic,

$$\int_{|y(t)|}^{\infty} p\left[\epsilon_y(t)\right] d\epsilon_y(t)$$

for floating-point arithmetic. Similarly, the conditional probability that $y(t)$ rounds to zero is

$$\int_{|y(t)| - 1/2}^{y(t) + 1/2} p\left[\epsilon_y(t)\right] d\epsilon_y(t)$$



(a) Unscaled Data (b) Data Scaled by 16 (c) Floating-Point Data

FIGURE III-18

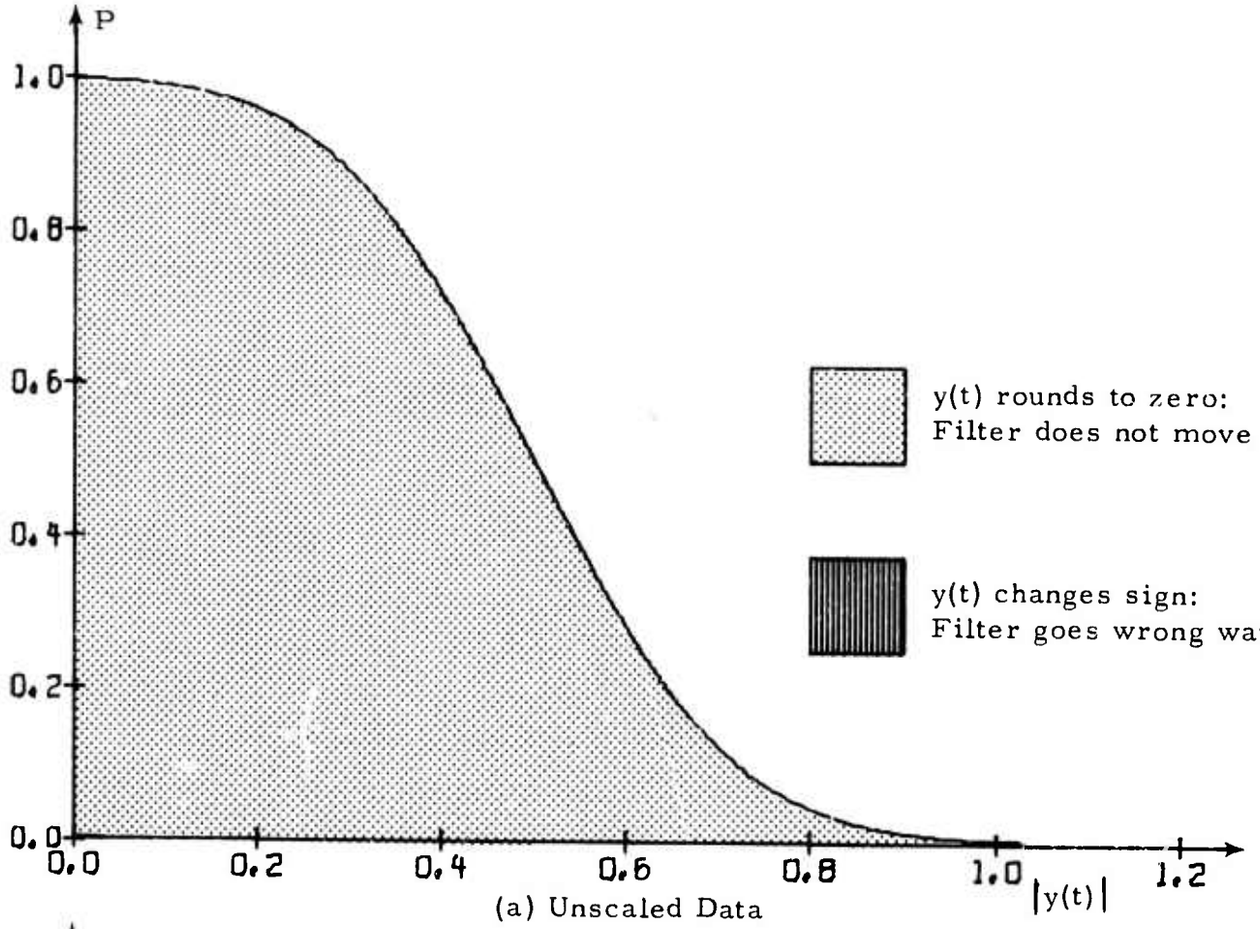
ADAPTIVE FILTER OUTPUT ERROR PROBABILITY DENSITY
PER COUNT BEFORE ROUNDOFF (MAXIMUM POSSIBLE VARIANCE)

for integer arithmetic, zero for floating-point arithmetic. The conditional probability that the sign of $y(t)$ is reversed or that $y(t)$ rounds to zero is displayed in Figure III-19 in the most favorable case for unscaled data, data scaled by 16, and floating-point data. The stippled areas designate the probability that $y(t)$ rounds to zero, the striped areas the probability that $y(t)$ changes sign. The corresponding probabilities in the most unfavorable case are depicted in Figure III-20.

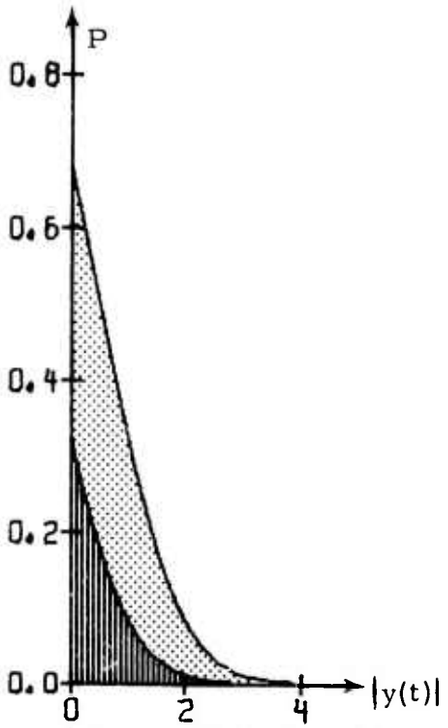
For integer arithmetic, the total probability that the sign of $y(t)$ changes is the integral

$$\int_0^{\infty} \left[\int_{|y(t)| + 1/2}^{\infty} p\left[\epsilon_y(t)\right] d\,\epsilon_y(t) \right] p\left[|y(t)|\right] d\,|y(t)|$$

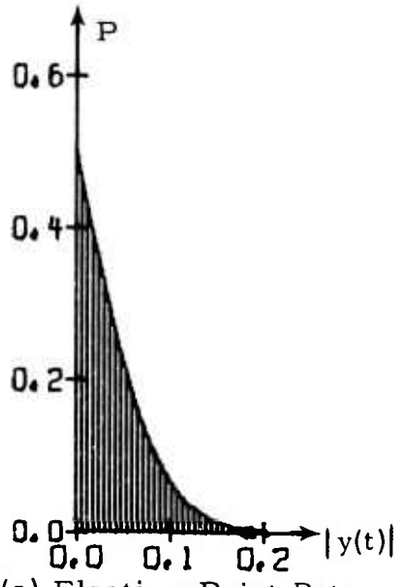
over the interval $(0, \infty)$ of the conditional probability of sign inversion given $|y(t)|$ multiplied by the probability density for the absolute value $|y(t)|$. The total probability that $y(t)$ rounds to zero is obtained by replacing the inner integral by the conditional probability that $y(t)$ is zero after roundoff: the interval of integration is from $|y(t)| - 1/2$ to $|y(t)| + 1/2$. For floating-point arithmetic, similarly, the limits of integration change to $|y(t)|$ and infinity when specifying the total probability of sign inversion. Figure III-21 illustrates the regions in the $\left[y(t), \epsilon_y(t)\right]$ plane corresponding to the total probabilities stipulated. To calculate these integrals, a knowledge of the probability distribution for the adaptive-filter output values is required. Unfortunately, it is not available. However, the adaptive-filter RMS level is available for a number of data samples. Under the assumption that $y(t)$ and the error $\epsilon_y(t)$ before roundoff are normally-distributed, independent random variables with zero mean, the total probabilities can be found. For integer arithmetic, the total probability of sign inversion is



(a) Unscaled Data



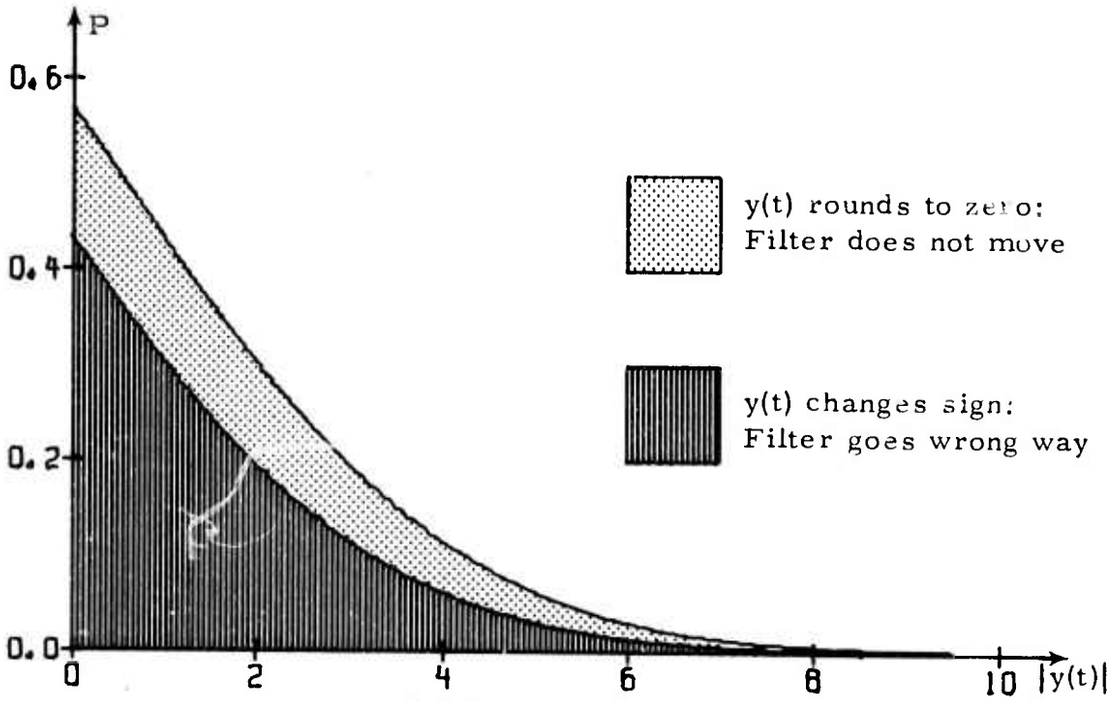
(b) Data Scaled by 16



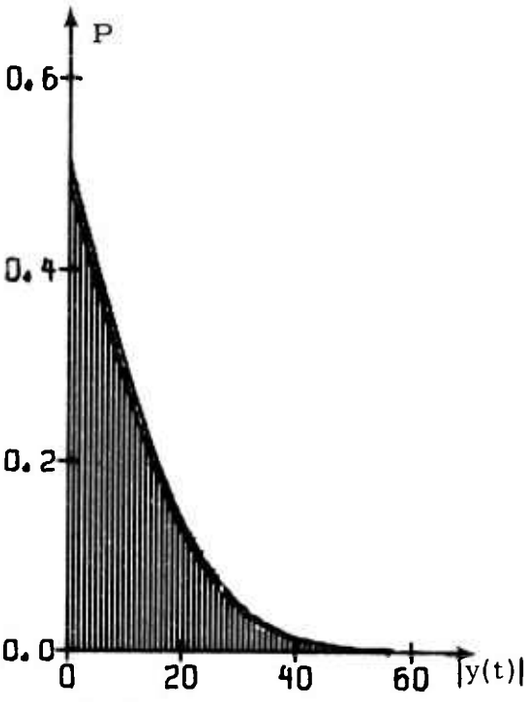
(c) Floating-Point Data

FIGURE III-19

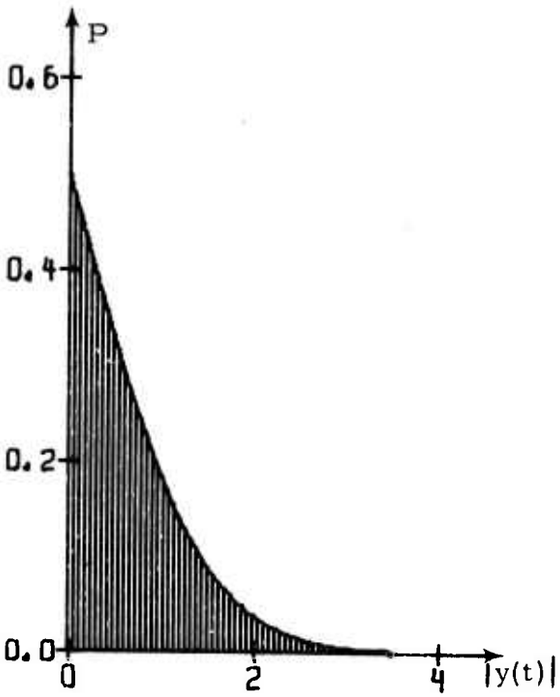
PROBABILITY THAT y(t) CHANGES SIGN OR ROUNDS TO ZERO
GIVEN |y(t)| [MOST FAVORABLE CASE: $y(t) = \bar{x}(t)$]



(a) Unscaled Data



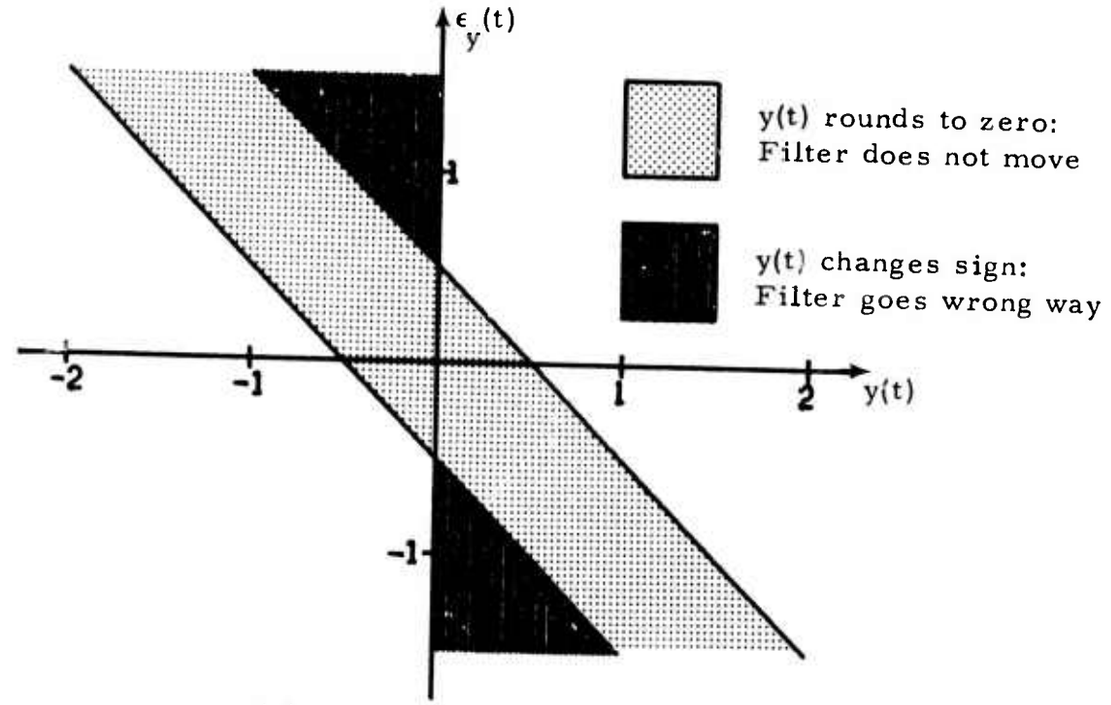
(b) Data Scaled by 16



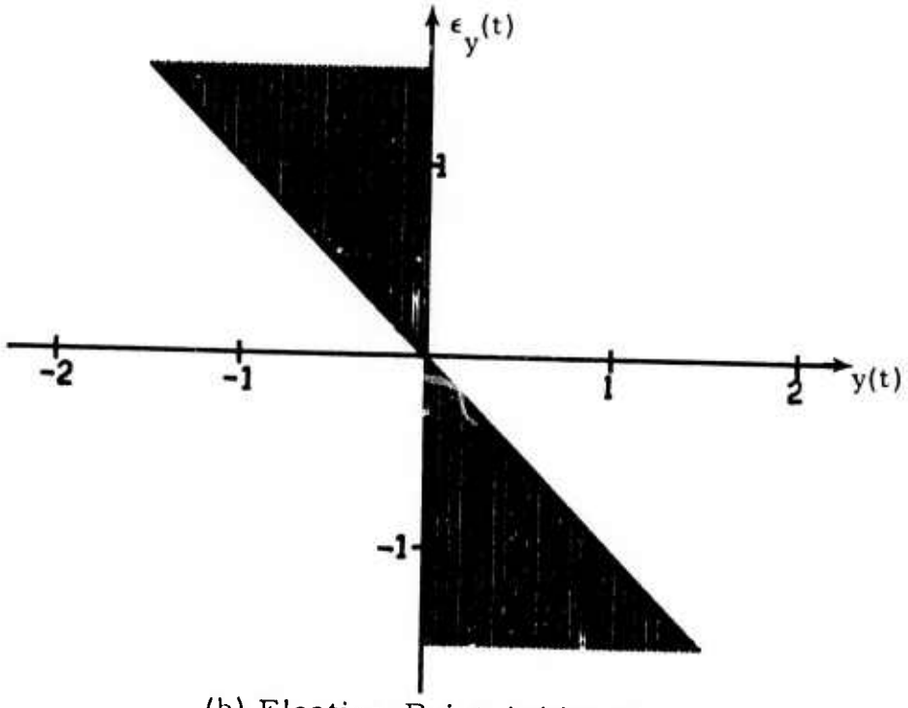
(c) Floating-Point Data

FIGURE III-20

PROBABILITY THAT y(t) CHANGES SIGN OR ROUNDS TO ZERO GIVEN $|y(t)|$ [MOST UNFAVORABLE CASE: $|a_j(j)| = 1/2$]



(a) Integer Arithmetic



(b) Floating-Point Arithmetic

FIGURE III-21

REGIONS IN THE $[y(t), \epsilon_y(t)]$ PLANE CORRESPONDING TO SIGN REVERSAL AND ROUNDING TO ZERO

$$\int_0^{\infty} \left[2 \int_{y+1/2}^{\infty} \frac{1}{\sigma_\epsilon \sqrt{2\pi}} \; e^{-(\epsilon_y^2 / 2\sigma_\epsilon^2)} \, d\epsilon_y \right] \frac{1}{\sigma_y \sqrt{2\pi}} \; e^{-(y^2/2\sigma_y^2)} \, dy$$

$$= 2 \int_0^{\infty} F\left(\frac{-1/2 - y}{\sigma_\epsilon}\right) \frac{e^{-(y^2/2\sigma_y^2)}}{\sigma_y \sqrt{2\pi}} \, dy \; ,$$

where σ_y is the adaptive-filter output RMS level and $F(z)$ is the standard cumulative normal distribution function

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \; e^{-z^2/2} \, dz \; .$$

The probability of rounding to zero is

$$F\left(\frac{1}{2\sqrt{\sigma_y^2 + \sigma_\epsilon^2}}\right) - F\left(-\frac{1}{2\sqrt{\sigma_y^2 + \sigma_\epsilon^2}}\right) \; .$$

For floating-point arithmetic, the probability that the sign of $y(t)$ changes is equal to the probability that the random variable

$$w = \frac{y(t)}{\epsilon_y(t)}$$

lies within the interval $-1 \leq w \leq 0$. The random variable w has a Cauchy distribution with probability density function

$$p(w) = \frac{\rho}{\pi(1 + \rho^2 w^2)} \; ,$$

where ρ is the ratio σ_ϵ/σ_y of the standard deviation of $\epsilon_y(t)$ to the RMS level of the adaptive filter output $y(t)$. The probability of a difference in sign between $y(t)$ and $y(t) + \epsilon_y(t)$, therefore, is

$$\frac{\rho}{\pi} \int_{-1}^{0} \frac{dw}{1 + \rho^2 w^2} = \frac{\tan^{-1}\rho}{\pi} .$$

The arctangent is the angle from the $(\epsilon_y/\sigma_\epsilon)$-axis to the line $\rho(\epsilon_y/\sigma_\epsilon) = -(y/\sigma_y)$ after the probability density is circularized by a coordinate transformation which divides each random variable by its standard deviation. This last estimate is useful even in integer arithmetic situations where the error $\epsilon_y(t)$ before roundoff has a standard deviation much larger than one half (e.g., Figure III-20b). Table III-4 lists the probability of a sign change or zero value in the filter output from a 270^o adaptive beam (with $K_s = 0.005$) for a four-hour noise sample from day 238 of 1970 under the assumption that all filter weights are of magnitude one half. These probabilities are uncomfortably high. They indicate that sign changes and zero values may significantly limit adaptive-filter performance. Even in the case of floating-point data, digitization error results in a surprisingly high probability of a filter update vector $A^{new} - A^{old}$ pointing in the wrong direction. The gain doubling performed on the ALPA seismometers in July and August of 1972 should reduce this probability by a factor of approximately two.

These appreciable errors in $y(t)$, particularly in the case of unscaled data, may account for the radically different noise reduction values at data scale factors of 1 and 16, respectively, for the quiet summer noise sample. At the convergence rate $K_s = 0.30$, the adaptive filter takes advantage of transient correlation between successive data vectors to reduce the filter output power. In order to do so effectively, the adaptive filter must react rapidly and accurately as each new data vector X is received. If the filter step

TABLE III-4

MAXIMUM PROBABILITY OF SIGN INVERSION OR ROUNDING TO ZERO IN ADAPTIVE FILTER OUTPUT (DAY 238, STEER DIRECTION 270°, $K_s = 0.005$)

	Standard Deviation σ_ϵ (counts)	RMS Level σ_y (counts)	Probability of Rounding To Zero	Probability of Sign Change
Unscaled Data	2.902	3.129	0.093	0.194
Data Scaled by 16	18.016	50.064	0.008	0.106
Floating-Point Data	1.113	3.129	0.000	0.109

$A^{new} - A^{old}$ pointed in the wrong direction a significant fraction of the time, the drop in noise reduction from 6 dB to 2 dB when a data scale factor of 1 was used instead of 16 could easily be explained. In the event that the adaptive filter output $y(t)$ rounded to zero, the effect would be somewhat less severe: the filter vector would not update and could not react to incoming data until the next data vector was available. Each time no update was performed, some of the ability to exploit the transient correlation between consecutive data vectors would be lost.

Without knowledge of the crosscorrelation statistics between input data channels, the unlikely possibility that all filter weights are of magnitude one half cannot be completely excluded. Part 2 of this subsection, as a by-product of other investigations, gives the filter weights obtained before and after the error vector E_c is added to the data vector X. For specific cross-correlation matrices, the absolute values of the resulting filter weights could be used to provide more accurate estimates of the error $\epsilon_y(t)$ before round-off in the adaptive filter output.

2. Change in the Mean Square Filter Output Power $\overline{y^2(t)}$ Due to Adding the Error Vector E_c to the Data Vector X

If the error vector E_c is added to the data vector X and the round-off error $\epsilon_r(t)$ after forming the dot product $(X + E_c)^T A$ is incorporated into the adaptive-filter output, a maximum-likelihood filter set can be designed to minimize the total output power subject to the maximum-likelihood constraints and an additional special constraint which reflects the fact that the term $\epsilon_r(t)$ is preserved with unity response. The adaptive-filter output with roundoff error is

$$y'(t) = (Q')^T A' = \left[(X + E_c)^T \vdots \epsilon_r(t)\right] \left[\begin{matrix} A \\ \cdots \\ b \end{matrix}\right] = \left[Q^T \vdots \epsilon_r(t)\right] \left[\begin{matrix} A \\ \cdots \\ b \end{matrix}\right]$$

where Q' has $1 + M(2N+1)$ components, A is subject to the normal maximum

likelihood constraints and the last component b of the expanded filter vector A' is subject to the special constraint b = 1. The constraints are expressed in the matrix equation

$$\left[\begin{array}{cccc|c} [1\cdots1] & [0\cdots0] & \cdots & [0\cdots0] & 0 \\ [0\cdots0] & [1\cdots1] & \cdots & [0\cdots0] & 0 \\ \vdots & & & & \vdots \\ [0\cdots0] & [0\cdots0] & [1\cdots1] & [0\cdots0] & 0 \\ [0\cdots0] & [0\cdots0] & [0\cdots0] & [1\cdots1] & 0 \\ \hline [0\cdots0] & [0\cdots0] & \cdots & [0\cdots0] & 1 \end{array}\right] \left[\begin{array}{c} \\ \\ \\ \\ \\ \hline b \end{array}\right] = \left[\begin{array}{c} \\ \\ \\ \\ \\ \hline 1 \end{array}\right] ,$$

where D is the (2N + 1)-dimensional vector

$$\left[\begin{array}{c} d(-N) \\ \vdots \\ d(-1) \\ d(0) \\ d(1) \\ \vdots \\ d(N) \end{array}\right] = \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{array}\right]$$

and each of the row subvectors $[1\cdots\cdots1]$ or $[0\cdots\cdots0]$ in the constraint matrix has M components (where M is the number of channels). In matrix form, it may be written

$$C'A' = D' ,$$

where the ordinary maximum-likelihood constraint matrix C is formed by dropping the last row and column from C', and where the expanded response vector D' is

$$D' = \begin{bmatrix} D \\ \hline 1 \end{bmatrix} .$$

C' is a $(2N + 2)$ by $[1 + M(2N + 1)]$ matrix, A' a $[1 + M(2N + 1)]$ - dimensional column vector reflecting the white frequency response of the filter A and the unalterable response of one on the error $\epsilon_r(t)$ generated by rounding the adaptive filter output to the nearest integer. The adaptive-filter updata equation uses the method of steepest descent to reduce the mean square output power $\overline{[y'(t)]^2}$ using the Widrow approximation $\overline{[y'(t)]^2} = [y'(t)]^2$ (Widrow, 1966):

$$(A')^{new} = A' - \mu \nabla \left\{ \left[y'(t) \right]^2 - 2 \left[(\Lambda')^T (D' - C'A') \right] \right.$$

$$= A' - 2\mu \left[y'(t)\, Q' + (C')^T \Lambda' \right] ,$$

where Λ' is the Lagrangian multiplier vector

$$\Lambda' = \begin{bmatrix} \lambda(-N) \\ \vdots \\ \lambda(-1) \\ \lambda(0) \\ \lambda(1) \\ \vdots \\ \lambda(N) \\ \hline \lambda_\epsilon \end{bmatrix} .$$

The Lagrangian multiplier vector is found by solving the vector equation

$$D' = C'(A')^{new} = C'A' - 2\mu \left[y'(t)\, C'Q' + C'(C')^T \Lambda' \right] :$$

$$\Lambda' = \left\{ \frac{1}{2\mu} \left[C'(C')^T \right]^{-1} \left[C'A' - D' \right] \right\} - y'(t) \left[C'(C')^T \right]^{-1} C'Q'$$

$$= -y'(t) \left[C'(C')^T \right]^{-1} C'Q' \ .$$

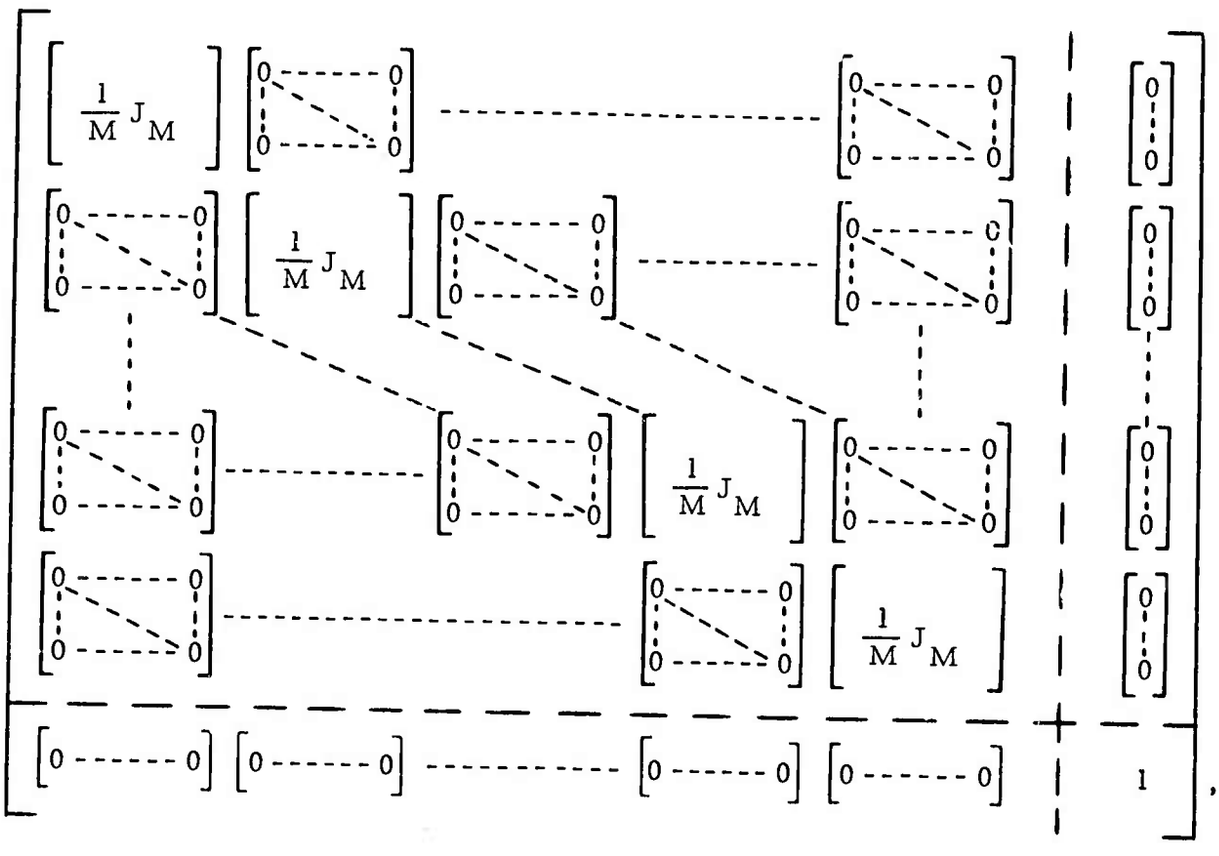
Substitution of Λ' into the update equation yields

$$(A')^{new} = A' - 2\mu\, y'(t) \left\{ I - (C')^T \left[C' (C')^T \right]^{-1} C' \right\} Q' \ .$$

The matrix $C'(C')^T$ is the $(2N + 2)$ by $(2N + 2)$ diagonal matrix

$$\left[\begin{array}{cccc|c} M & 0 & \cdots & 0 & 0 \\ 0 & M & 0 \cdots & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & M \quad 0 & 0 \\ 0 & \cdots & 0 & 0 \quad M & 0 \\ \hline 0 & 0 & \cdots & 0 \quad 0 & 1 \end{array}\right]$$

and the matrix $(C')^T \left[C' (C')^T \right]^{-1} C'$ is the $[1 + M(2N + 1)]$ by $[1 + M(2N + 1)]$ matrix



where each of the $2N+1$ submatrices J_M is an M by M matrix with each element equal to 1. The update equation reduces to

$$(A')^{new} = A' + 2\mu y'(t)\,(\overline{Q}' - Q') \quad ,$$

where $\overline{Q'} = (C')^T \left[C'(C')^T\right]^{-1} C'Q'$, or the equivalent equation

$$\begin{bmatrix} A^{new} \\ \hline 1 \end{bmatrix} = \begin{bmatrix} A \\ \hline 1 \end{bmatrix} + 2\mu y'(t) \begin{bmatrix} \overline{Q} - Q \\ \hline 0 \end{bmatrix}$$

where $\overline{Q} = C^T \left[CC^T\right]^{-1} CQ$. This equation, which is the equation for minimizing the total output power $\overline{\left[y'(t)\right]^2}$ is precisely the equation which is implemented when the error vector E_c is added to the data vector X if no roundoff error occurs when the beamsteer output components

$$\bar{q}_i(t-j) = \frac{1}{M} \sum_{i=1}^{M} q_i(t-j)$$

are rounded to the nearest integer:

$$A^{new} = A + 2\mu (Q')^T A' \left[\bar{Q}-Q\right]$$

$$= A + 2\mu \left[(X + E_c)^T A + \epsilon_r(t) \right] \left[\overline{(X + E_c)} - (X + E_c) \right].$$

As was shown in Subsection C, the beamsteer roundoff error vector E_b can be eliminated by scaling the vector Q by the number of channels and dividing the convergence factor μ by M:

$$M\bar{q}_i(t-j) = \sum_{i=1}^{M} q_i(t-j) ,$$

so that $M\bar{q}_i(t-j)$ is always an integer and no roundoff error occurs. By choosing μ small enough, the adaptive filter set can be shown to converge to the optimum maximum-likelihood filter set if certain conditions are met (Daniell, 1968).

This relationship between the adaptive filter set and the optimum maximum-likelihood filter set can be used to calculate the mean square filter output power $\overline{[y'(t)]^2}$ with digitization and roundoff error. $\overline{[y'(t)]^2}$ can be compared with the mean square filter output power $\overline{y^2(t)}$ in the absence of such errors. The adaptive filter output with these errors is

$$y'(t) = y(t) + \epsilon_y(t) + \epsilon_r(t) ,$$

where $\epsilon_y(t)$ is the error before roundoff and $\epsilon_r(t)$ is the roundoff error occurring when $y(t) + \epsilon_y(t)$ is rounded to the nearest integer. The quantity

$y(t) + \epsilon_y(t)$ may be broken up into the integer

$$\left[y(t) + \epsilon_y(t)\right]_I$$

obtained after roundoff and the fractional term

$$\left[y(t) + \epsilon_y(t)\right]_F$$

which is eliminated by the addition of $\epsilon_r(t)$. The mean square noise power $\overline{[y'(t)]^2}$ is

$$\overline{[y'(t)]^2} = \overline{[y(t) + \epsilon_y(t)]^2} + 2\,\overline{[y(t) + \epsilon_y(t)]_I\ \epsilon_r(t)}$$

$$+ 2\,\overline{[y(t) + \epsilon_y(t)]_F\ \epsilon_r(t)} + \overline{\epsilon_r^2(t)} \quad .$$

When the standard deviation of $y(t) + \epsilon_y(t)$ is larger than one count, its rounded value $[y(t) + \epsilon_y(t)]_I$ has only negligible correlation with the round-off error $\epsilon_r(t)$, so that the term

$$2\,\overline{[y(t) + \epsilon_y(t)]_I\ \epsilon_r(t)}$$

can be neglected. On the other hand,

$$[y(t) + \epsilon_y(t)]_F + \epsilon_r(t) = 0 ,$$

so that

$$\overline{[y'(t)]^2} = \overline{[y(t) + \epsilon_y(t)]^2} - \overline{2\,\epsilon_r^2(t)} + \overline{\epsilon_r^2(t)}$$

$$= \overline{\left[y(t) + \epsilon_y(t)\right]^2} - \overline{\epsilon_r^2(t)}$$

$$= A^T \overline{(X + E_c)(X + E_c)^T} A - \overline{\epsilon_r^2(t)}$$

$$= A^T \Phi A - \overline{\epsilon_r^2(t)} \ ,$$

where the matrix Φ is the crosscorrelation matrix corresponding to the data vectors $X + E_c$. Minimizing the mean square filter output power $\overline{[y'(t)]^2}$ is the same as minimizing the term $A^T \Phi A$.

When the filter weights are subject to the maximum-likelihood constraints $CA = D$, the optimum filter set is achieved when each component of the gradient

$$\nabla \left[1/2 \ A^T \Phi A + \Lambda^T (D - CA)\right] = \Phi A - C^T \Lambda$$

is zero. The Lagrangian multiplier vector Λ is found by using the constraint conditions $CA = D$, so that the matrix equation

$$\begin{bmatrix} \Phi & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} A \\ -\Lambda \end{bmatrix} = \begin{bmatrix} 0 \\ D \end{bmatrix}$$

yields a solution for both the filter vector A and the Lagrangian multiplier vector Λ. By rearranging the rows and columns of this matrix equation, the matrix can be put in block-Töplitz form (as on page II-3). The filter

output power $A^T\Phi A$ is the zero-lag component $\lambda(0)$ of the Lagrangian multiplier vector:

$$\Phi A = C^T \Lambda$$

$$A = \Phi^{-1} C^T \Lambda$$

$$CA = (C\Phi^{-1}C^T)\Lambda = D$$

$$\Lambda = (C\Phi^{-1}C^T)^{-1} D$$

$$A = \Phi^{-1} C^T (C\Phi^{-1}C^T)^{-1} D$$

$$A^T\Phi A = D^T (C\Phi^{-1}C^T)^{-1} C\Phi^{-1}\Phi\,\Phi^{-1}C^T(C\Phi^{-1}C^T)^{-1} D$$

$$= D^T (C\Phi^{-1}C^T)^{-1} D$$

$$= D^T \Lambda$$

$$= \lambda(0) \ .$$

The vectors Λ and A as well as the mean square noise power $A^T\Phi A$ are immediately available upon solution of the full matrix equation. The filter output power with the addition of the error vector E_c and the roundoff error $\epsilon_r(t)$ is

$$D^T (C\Phi_{X+E} C^T)^{-1} D - \overline{\epsilon_r^2(t)} = \lambda_{X+E}(0) - \overline{\epsilon_r^2(t)} \ ,$$

where the subscript $X+E$ denotes the addition of the error vector E_c to the data vector X. Without digitization error and roundoff error, the filter output power is

$$D^T (C\Phi_X C^T)^{-1} D = \lambda_X(0) \ ,$$

where the subscript X denotes the data vector X by itself. The ratio of the altered mean square filter output power $\overline{[y'(t)]^2}$ to the original mean square filter output power $\overline{y^2(t)}$ is

$$\frac{\lambda_{X+E}(0) - \overline{\epsilon_r^2(t)}}{\lambda_X(0)} \quad .$$

Provided that the mean square error in predicting one channel from any other channel at any time shift up to $N\Delta t$ is larger than 1 count^2, a channel value rounded to the nearest integer (and the roundoff error on that channel) should have negligible correlation with the roundoff error on any other channel. As a result, crosscorrelation values between two different channels should be unaffected by digitization and roundoff error. Autocorrelation values, however, are affected. Under reasonably weak assumptions concerning the lack of perfect correlation between triax components at each site, between successive values of the unfiltered vertical component at each site, and between successive values of the prefiltered vertical component at each site, the autocorrelation values after allowing for digitization error and roundoff error should be

$$\overline{x_i'(t)\, x_i'(t+\tau)} = \overline{x_i(t)\, x_i(t+\tau)} - \left[\frac{1}{3}\,\overline{\epsilon_d^2} + \overline{\epsilon_v^2}\right]\varphi_b(\tau) - \delta_{0\tau}\,\overline{\epsilon_x^2} \quad ,$$

where the primes indicate values with error, where $\overline{\epsilon_d^2}$ is the variance of the digitization error on each triax sensor, where $\overline{\epsilon_v^2}$ is the variance of the error in rounding each unfiltered vertical component to the nearest integer, where $\overline{\epsilon_x^2}$ is the variance of the error in rounding off each prefiltered vertical component, where $\varphi_b(\tau)$ is the autocorrelation function

$$\varphi_b(\tau) = \sum_{k=-L}^{L} b_k b_{k+\tau}$$

of the prefilter applied to the unfiltered vertical-component data, and where $\delta_{o\tau}$ is the Kronecker delta operator (the variance $\overline{\epsilon_x^2}$ is subtracted only when $\tau = 0$). Floating-point autocorrelation functions are affected only by digitization error:

$$\overline{x_i'(t)\, x_i'(t+\tau)} = \overline{x_i(t)\, x_i(t+\tau)} - \frac{1}{3}\, \overline{\epsilon_d^2}\, \varphi_b(\tau) \quad .$$

In the case of specific noise matrices Φ_X , these results provide a convenient technique for determining the change in the mean square noise power $\overline{y^2(t)}$ from errors in the data vector X and from roundoff error after the dot product $(X + E_c)^T A$ has been formed. A knowledge of the noise environment in which an adaptive-filtering system is to operate, together with a computation of the digitization and roundoff error statistics associated with a particular implementation scheme, is sufficient to set specifications on the A/D system and on the precision of the data vector X so as to come arbitrarily close to the mean square noise level $\lambda_X(0)$ for data without digitization or roundoff error. (Further specifications on the precision of the filter vector A are, of course, also necessary.)

In one data sample with an enormous signal, it was necessary to reduce the data scale factor from 16 to 4 to avoid clipping the signal. After three runs were made at scale factors of 16, 8, and 4, the noise reduction values over the section of data preceding the signal were sufficiently interesting to make two additional runs at scale factors of 2 and 1. The noise reduction results in Table III-5 are for the period 0430 to 0638 on day 7 of 1972. An adaptive filter was steered toward 253^o at a convergence rate $K_s = 0.005$. Decibel values in this table are relative to 1 count^2/second. With the exception of the varying data scale factors, all processing parameters are as described in Subsection A of Section V. Note that the beam-steer output power increases by 6.181 dB, 6.098 dB, 6.061 dB, and 6.041 dB,

TABLE III-5

NOISE REDUCTION AS A FUNCTION OF DATA SCALE FACTOR
(DAY 7 1972, 0430 to 0648)

	Beamsteer Output Power (dB)	Adaptive Filter Output Power (dB)	Noise Reduction (dB)
Unscaled Data	18.844	16.540	2.304
Data Scaled by 2	25.025	22.780	2.245
Data Scaled by 4	31.123	28.906	2.217
Data Scaled by 8	37.184	34.986	2.198
Data Scaled by 16	43.225	41.037	2.188

respectively, each time the data scale factor is doubled. Similarly, the adaptive filter output power increases by 6.260 dB, 6.126 dB, 6.080 dB, and 6.051 dB. Without roundoff error, these figures should increase by $20 \log_{10} 2$ (6.021 dB). The actual results reflect the fact that roundoff error reduces the mean square noise power by a relatively stable number of squared counts. The only significance of these results is that zero output is obtained by scaling the data down to the point where the largest value $x_i(t-j)$ in the data lies between -1/2 and 1/2 before it is rounded off. The same effect occurs when the crosscorrelation matrix is formed from data containing a signal. In fact, signal power might be reduced more than noise power. Crosscorrelation terms as well as autocorrelation terms in an ideal signal matrix are affected:

$$\overline{s_i'(t)\, s_k'(t+\tau)} = \overline{s_i'(t)\, s_i'(t+\tau)}$$

$$= \overline{s_i(t)\, s_i(t+\tau)} - \left[\frac{1}{3} \overline{\epsilon_d^2} + \overline{\epsilon_v^2} \right] \varphi_b(\tau) - \delta_{o\tau} \overline{\epsilon_x^2} \; .$$

As the data is scaled down to the point where the mean square error in predicting one channel from another drops well below 1 count^2, the crosscorrelation values need adjustment. And this condition is more readily achieved when a signal is present.

One important fact can be inferred from this study of the effects of error in the data vector X. It is that the principal effect of roundoff error and digitization error at very low convergence rates is simply to reduce the autocorrelation function values in the matrix $\Phi = \overline{XX^T}$ when a significant amount of power (in terms of computer counts) is uncorrelated between sensors. Normally, the reduction of the autocorrelation functions causes a drop in the filter output power. Aside from this scaling effect on signal and noise alike, the error appears at very low convergence rates to have no major consequences in terms of the mean square filter output once spatially uncorrelated

noise rises to certain minimal computer-count levels. This statement applies, of course, only to situations where the convergence rate is slow enough that the adaptive filter set closely approximates the optimum maximum-likelihood filter set. It does not apply to the effect of digitization and roundoff error on misadjustment, which is defined to be the fraction

$$\frac{\overline{y^2(t)} - \left[\overline{y^2(t)}\right]_{\text{optimum}}}{\left[\overline{y^2(t)}\right]_{\text{optimum}}}$$

of additional noise caused by the adaptive algorithm in a time-stationary environment. The drop in noise reduction from 6 dB to 2 dB at $K_s = 0.30$ clearly illustrates this fact in the case of the quiet summer noise sample where data scale factors of 16 and 1 were used.

An appropriate design goal would be to reduce the difference between the mean square noise output values with and without error to the maximum extent feasible. In that event, at least, processing results would increasingly reflect instrument measurements rather than computational noise.

E. DIRECTIONAL ERROR OF THE FILTER UPDATE VECTOR $(A^{new} - A^{old})$

In this subsection, the effects of rounding off the filter update vector are investigated. Part 1 discusses only the roundoff problems involved in the direct calculation of the update vector. After this calculation, the constraint conditions may no longer be satisfied. As has been pointed out in the literature (Frost, 1972), the cumulative effect of repeated calculations not satisfying the constraint conditions is to produce a random walk away from the constraint space. Some corrective procedure must be performed periodically to bring the filter set back into reasonable agreement with the constraint conditions. Part 2 considers the implications of various corrective methods on

the directional error of the filter update vector. Part 3 briefly describes the disadvantages of a procedure which eliminates filter-weight roundoff error and beamsteer roundoff error.

1. Error Due to Rounding Off the Filter Weights $a_i(j)$

The average change in an individual component of the filter weight vector A is a critical quantity in evaluating the error due to rounding off the filter weights. An individual component may be immobilized or partially immobilized when the average change drops significantly below the quantization level of the filter weights. Even when the average change is approximately the same as the filter weight quantization level, the fact that some filter weight components $a_i(j)$ will be hampered more than others in their movement means that the filter vector A may be deflected from the desired direction of movement after several iterations.

The squared magnitude $|A^{new} - A^{old}|^2$ of the filter update vector is

$$(A^{new} - A^{old})^T \; (A^{new} - A^{old})$$

$$= \frac{4K_s^2 \; A^T X X^T A \; (\bar{X}-X)^T \; (\bar{X}-X)}{(\bar{X}-X)^T \; (\bar{X}-X) \; (\bar{X}-X)^T \; (\bar{X}-X)}$$

$$= \frac{4K_s^2 \; y^2(t)}{(\bar{X}-X)^T \; (\bar{X}-X)}$$

for the particular adaptive algorithm used in this report. With some mathematical manipulations, a clearer picture emerges of the factors influencing the filter step size:

$$\left|A^{new} - A^{old}\right|^2 = \frac{4K_s^2\, y^2(t)}{(\bar{X}-X)^T(\bar{X}-X)}$$

$$= \frac{4K_s^2\, \Psi^2(t)\, E\left[y^2(t)\right]}{E\left[(\bar{X}-X)^T(\bar{X}-X)\right]}$$

$$= \frac{4K_s^2\, \Psi^2(t)}{\dfrac{E\left[\bar{X}^T\bar{X}\right]}{E\left[y^2(t)\right]} \quad \dfrac{E\left[(\bar{X}-X)^T(\bar{X}-X)\right]}{E\left[\bar{X}^T\bar{X}\right]}}$$

$$= \frac{4K_s^2\, \Psi^2(t)}{\dfrac{M(2N+1)\, E\left[\bar{x}^2(t)\right]}{E\left[y^2(t)\right]} \quad \dfrac{E\left[X^TX\right] - E\left[\bar{X}^T\bar{X}\right]}{E\left[\bar{X}^T\bar{X}\right]}}$$

$$= \frac{4K_s^2\, \Psi^2(t)}{M(2N+1)\, R_A^2 (R_B^2 - 1)} \quad ,$$

where the operator $E[---]$ denotes the time average of the quantity inside the brackets, where $\Psi(t)$ is the random variable

$$\Psi(t) = \sqrt{\frac{y^2(t)}{E\left[y^2(t)\right]} \Big/ \frac{(\bar{X}-X)^T(\bar{X}-X)}{E\left[(\bar{X}-X)^T(\bar{X}-X)\right]}} \quad ,$$

where R_A^2 is the beam output power reduction

$$R_A^2 = \frac{E\left[\bar{x}^2(t)\right]}{E\left[y^2(t)\right]}$$

of adaptive filtering relative to beamsteering, and where R_B^2 is the power reduction

$$R_B^2 = \frac{E\left[X^T X\right]}{E\left[\bar{X}^T \bar{X}\right]}$$

of beamsteering relative to the average single-sensor power. R_B^2 is normally close to the value M (where M is the number of channels) provided that the noise field is spatially uncorrelated or coherent but distributed over a range of look directions not impinging on the main lobe of the time-shift-and-sum array beam pattern. Under the assumption that $R_B^2 = M$, the RMS change in one component of the filter weight vector is

$$\sqrt{\frac{\left|A^{new} - A^{old}\right|^2}{M(2N+1)}} = \frac{2K_s \Psi(t)}{M(2N+1)\sqrt{M-1}\ R_A} .$$

It is the RMS change in the sense that it is the square root of the squared components $a_i^2(j)$ averaged over all components. The quantity in question still fluctuates as a function of time. If the time averages $E[---]$ are taken over time intervals in which the power reduction R_A^2 of adaptive filtering relative to beamsteering is stable, the RMS change in a filter component is the product of the stable value

$$\frac{2K_s}{M(2N+1)\sqrt{M-1}\ R_A}$$

and the time-varying random variable $\Psi(t)$. In view of the definition of $\Psi(t)$,

it is likely that the mode point of the probability distribution for $\Psi(t)$ is close to 1. As a result, the mode point of the probability distribution for the RMS change in one filter component $a_i(j)$ is close to

$$\frac{2K_s}{M(2N+1)\sqrt{M-1}\ R_A} .$$

At the convergence rate $K_s = 0.005$ and the amplitude reduction factor $R_A = 1.333$ (about 2.5 dB) for adaptive filtering relative to beamsteering, this quantity is 1/65536, or 2^{-16}. The quantization level of the filter weights for the adaptive processor used in this report is 2^{-16}. (A filter weight of one half is represented as 32768 counts.) Thus the RMS filter-weight change was close to one count when the adaptive filter output power was 2.5 dB below the beamsteer output power. And the actual noise reduction obtained at a convergence rate $K_s = 0.005$ was typically below 2.5 dB, so that the mode point of the probability distribution for the RMS filter-weight jump was probably within the range 1 to 1.333 counts.

If the probability density function for the random variable $\Psi(t)$ were strongly peaked near a value of one, there would be considerable cause for alarm when the RMS change in one component of the filter weight vector dropped below 1/2 count. In that event, most of the components would be immobilized. A broadly dispersed probability distribution for $\Psi(t)$, on the other hand, would permit movement when $\Psi(t)$ reached its highest values (even if the rate K_s might be biased either upward or downward, but at least the filter vector would move. For this reason, the fluctuation of $\Psi(t)$ as well as the RMS change in $a_i(j)$ warrants scrutiny. The random variable $\Psi^2(t)$ may be expressed as the product

$$\Psi^2(t) = \left\{ \frac{y^2(t)}{E\left[y^2(t)\right]} \right\} \left\{ \frac{E\left[(\bar{X}-X)^T(\bar{X}-X)\right]}{(\bar{X}-X)^T(\bar{X}-X)} \right\} .$$

Most of the variation in $\Psi^2(t)$ can be attributed to the term on the left since the filter output $y(t)$ oscillates back and forth between positive and negative values, ordinarily reaching an amplitude greater than its RMS level. The term on the right, however, does not fluctuate nearly as much since it is a power average over $2N + 1$ consecutive points in time. The 31-second length of the adaptive filter used in this study spans 3 quarter-cycles of the 40-second period at which the bias-removal prefilter response (as well as the filtered data power spectrum) begins to roll off. Furthermore, power surges in $y^2(t)$ tend to compensate for power surges in $(\bar{X}-X)^T(\bar{X}-X)$. Since the point-to-point fluctuations in $y^2(t)$ have no counterpart in any other term affecting $\Psi^2(t)$, it is reasonably safe to assume that $\Psi^2(t)$ varies at least as much as $y^2(t)/E\left[y^2(t)\right]$. There are a number of probabilistic models that could conceivably describe the fluctuations in $y(t)$. Just which one is most appropriate depends on the data. A reasonable assumption for the sake of discussion is that $y(t)$ is normally distributed. In that event, it is easy to determine whether a "typical" filter-weight component is in danger of being immobilized by roundoff error when it is rounded to the nearest filter-weight count. By a "typical" component is meant a filter-weight component whose RMS change over time agrees with the RMS filter-weight change averaged across components. Table III-6 gives the probability of no movement and the standard deviation of the "typical" filter-weight component's motion after roundoff at four RMS filter-weight change values. In this table, only the variation in $\Psi^2(t)$ due to the term $y^2(t)/E\left[y^2(t)\right]$ has been considered, and the mean of the movement in the "typical" filter-weight component has been ignored. Under these assumptions, it is apparent that the "typical" component is severely impeded in its movement only when the RMS filter-weight change drops below 1/4 count.

TABLE III-6

PROBABILITY OF NO MOVEMENT AND STANDARD DEVIATION OF THE CHANGE IN A FILTER-WEIGHT COMPONENT WHOSE RMS CHANGE OVER TIME IS THE SAME AS THE RMS FILTER-WEIGHT CHANGE ACROSS COMPONENTS

RMS Filter-Weight Change (counts)	Probability of No Movement	Standard Deviation of the Change in a Typical Component (counts)
1.000	0.3830	1.04
0.500	0.6826	0.57
0.250	0.9546	0.21
0.125	1.0000	0.00

In reality, a wide variation in average movement can be expected across the components of the filter vector A. It is well within the realm of possibility that the bulk of the motion could be concentrated in a small number of components. At an RMS filter-weight change of 1/4 count, therefore, it is likely that most of the filter weights are greatly restricted in their activity. As a result, a safety margin of several bits in the numerical representation of the filter weights seems appropriate. On the other hand, it is true that those components most critical to adaptive-filter performance do move at the 1/4-count RMS change value. One unambiguous result can be derived from this analysis: the filter vector proper is unlikely to move at all when the RMS filter-weight change goes below $1/\left[8\sqrt{M(2N+1)}\right]$ counts; at that level, Table III-6 implies that the largest change in any component rounds to zero even if all of the change is concentrated in a single component. The empirical results of Subsection IV-B indicate that noticeable degradation in filter performance begins to occur when the RMS change in an individual filter weight drops below 1/2 count.

Due to the fact that the error in rounding off a filter weight never changes the sign of the filter weight, it is possible to place some limits on the angular error in the vector $(A^{new} - A^{old})$. It is never more than 90^{o} different from the vector $\mathrm{sgn}\left[y'(t)\right](\bar{X}-X + E_b + E_c)$, where sgn is the function

$$\mathrm{sgn}\, z = \begin{cases} -1 \text{ if } z < 0 \\ 0 \text{ if } z = 0 \\ 1 \text{ if } z > 0 \end{cases},$$

where $y'(t)$ is the adaptive filter output after roundoff, and where $\bar{X}-X$, E_b, and E_c are as defined previously. A slightly more useful bound can be found. Provided that the filter vector moves, the worst possible situation that can happen is that one component changes slightly more than 1/2 count while all other components change just slightly less than 1/2 count and round to zero.

In that event, the angle between $(A^{new} - A^{old})$ after roundoff and $\text{sgn}\left[y'(t)\right]\ (\bar{X}\text{-}X + E_b + E_c)$ is

$$\tan^{-1}\ \sqrt{M(2N+1)\ -1}\quad ,$$

or $85^{o}48'$ when $M(2N + 1) = 186$.

Up to this point, no accurate estimates of the angular error in the filter update vector $(A^{new} - A^{old})$ have been made because the change in the filter-weight components $a_i(j)$ before roundoff has been smaller than 1 count for at least some components in the situations examined. If all components change by at least several counts, simplifying assumptions which facilitate mathematical treatment are possible. In adaptive filtering systems where the process of rounding the filter weights affects the vector $(A^{new} - A^{old})$ only slightly, all or almost all of the filter-weight components do change by at least several counts. Hence it is worthwhile examining the results inferred from the simplifying assumptions.

The most easily derived result is an upper bound for the angle between the vector $\text{sgn}\left[y'(t)\right]\ (\bar{X}\text{-}X + E_b + E_c)$ and the vector $(A^{new} - A^{old})$ after roundoff. Prior to rounding the filter update vector to the nearest filter-weight count, the filter update equation is

$$\frac{A^{new} - A^{old}}{q} = \frac{2K_s\, y'(t)\ (\bar{X}\text{-}X + E_b + E_c)}{q\left|\ \bar{X}\text{-}X + E_b + E_c\right|^2}$$

$$= K\ (\bar{X}\text{-}X + E_b + E_c)\quad ,$$

where q is the quantization level of the filter weights (2^{-16} for the algorithm used in this report), and where K is the scalar constant

$$K = \frac{2K_s \; y'(t)}{q\left|\bar{X}-X + E_b + E_c\right|^2} \quad .$$

To the vector $(A^{new} - A^{old})/q$ is added the roundoff vector

$$D = \left[\begin{bmatrix}\delta_1 & (-N)\\ \vdots & \\ \delta_M & (-N)\end{bmatrix} \cdots \begin{bmatrix}\delta_1 & (0)\\ \vdots & \\ \delta_M & (0)\end{bmatrix} \cdots \begin{bmatrix}\delta_1 & (N)\\ \vdots & \\ \delta_M & (N)\end{bmatrix}\right] \quad .$$

The squared magnitude of the vector $(A^{new} - A^{old})/q$ is

$$\frac{4K_s^2 \left[y'(t)\right]^2}{q^2 \left|\bar{X}-X + E_b + E_c\right|^2}$$

in counts^2. The vector $(A^{new} - A^{old})/q$ <u>after roundoff</u> must lie within $|D|$ counts of $(A^{new} - A^{old})/q$ <u>before roundoff</u>. This situation is illustrated in

Figure III-22. The largest angle γ between $K(\bar{X}-X + E_b + E_c) + D$ and $K(\bar{X}-X + E_b + E_c)$ occurs when D is perpendicular to $K(\bar{X}-X + E_b + E_c)+ D$, so that

$$\sin^2 \gamma \leq \frac{q^2 |D|^2 |\bar{X}-X + E_b + E_c|^2}{4K_s^2 \left[y'(t)\right]^2}$$

Since the largest possible absolute value of any component $\delta_i(j)$ in D is 1/2, $|D|^2_{max} = M(2N + 1)/4$. Provided that the filter vector moves after roundoff, the minimum absolute value of $y'(t)$ is 1. Therefore

$$\sin^2 \gamma \leq \frac{46.5\, q^2 |\bar{X}-X + E_b + E_c|^2_{max}}{4\, K_s^2}$$

for a 186-component filter vector. In a four-hour noise sample from day 238 of 1970, the largest value $|\bar{X}-X + E_b + E_c|^2$ observed was 18,617,020 counts^2 when the data was scaled by a factor of 16. Under the assumption that the largest value $|\bar{X}-X + E_b + E_c|^2$ for unscaled data was 1/256 as large, Table III-7 lists the maximum possible angle between the filter update vector before roundoff and the filter update vector after roundoff at the convergence rate $K_s = 0.005$. This table indicates, in the case of unscaled data, that a 24-bit filter vector guarantees an angle γ considerably smaller than the average angle $\bar{\varphi}$ between the vectors $(\bar{X}-X)$ and $(\bar{X}-X + E_b + E_c)$ at the convergence rate $K_s = 0.005$. Twenty-eight bits accomplish the same purpose for data scaled by 16. When the data is scaled by 16, or course, it is far less likely that the rounded adaptive filter output $y'(t)$ has an absolute value of 1. The angular bounds in Table III-7 are extremely conservative. Whenever the maximum angle is below $85^o48'$, the following conditions must occur simultaneously to achieve it: (1) the squared magnitude $|\bar{X}-X + E_b + E_c|^2$

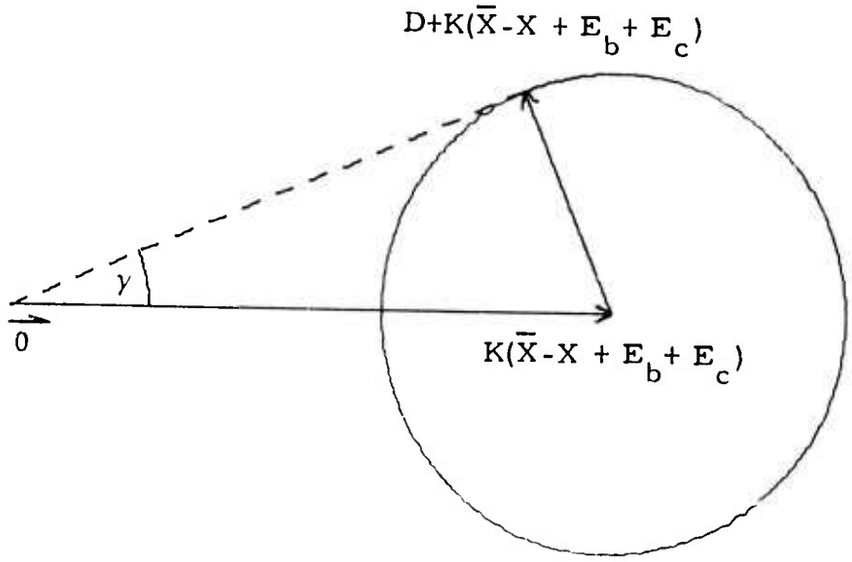


FIGURE III-22

TWO-DIMENSIONAL PLANE PASSING THROUGH THE ORIGIN, D, AND K($\bar{X}$-X + E_b+ E_c)

TABLE III-7

MAXIMUM POSSIBLE ANGLE OF DEFLECTION DUE TO ROUNDING THE FILTER UPDATE VECTOR AT THE CONVERGENCE RATE $K_s = 0.005$ (USING NOISE DATA FROM DAY 238 OF 1970)

$\text{Log}_2(1/q)$		
Unscaled Data	Data Scaled by 16	Maximum Angle
16	20	85°48'
17	21	85°48'
18	22	44°33'
19	23	20°32'
20	24	10° 6'
21	25	5° 2'
22	26	2°31'
23	27	1°15'
24	28	0°38'

must be at its maximum value; (2) the rectified filter output $|y'(t)|$ must be one; (3) the vector D must be perpendicular to the vector $K(\bar{X}-X+E_b+E_c)+D$; and (4) each component of D must be of magnitude 1/2.

The vector $(A^{new} - A^{old})$ without any digitization or roundoff error points in the same direction as the vector $\text{sgn}[y(t)](\bar{X}-X)$. With the errors incorporated in the vectors E_b, E_c, and D, it points in the same direction as the vector $D + |K| \text{ sgn } [y'(t)] (\bar{X}-X + E_b + E_c)$. Since the filter-weight roundoff vector D is symmetrically distributed about the origin and is zero when $y'(t) = 0$, the specification of the probability distribution for the angular error in $(A^{new}-A^{old})$ reduces to the determination of the probability distribution for the angle between $|K|$ $(\bar{X}-X)$ and $\text{sgn}[y(t)] \text{ sgn}[y'(t)][|K|(X-X + E_b + E_c) + D]$. Probability distributions for E_b and E_c were derived previously in Subsection C. Likewise, the probability distribution for $\text{sgn}[y(t)] \text{ sgn}[y'(t)]$ was discussed in Subsection D. As a first step in finding the probability distribution for the angle of error in $(A^{new}-A^{old})$ the filter-weight roundoff vector D needs to be examined.

Prior to rounding the filter weights at the j-th lag, an individual component of the filter vector is of the form

$$K\left\{\bar{x}(t-j) - x_i(t-j) + \epsilon_i(t-j) - \left[\frac{1}{M}\sum_{k=1}^{M} \epsilon_k(t-j)\right] + \epsilon_b(t-j)\right\} .$$

It consists of an integer part (the component rounded to the nearest integer) and a fractional part which is eliminated by the roundoff procedure. Thus, it is clear that the component $\delta_i(j)$ of the vector D is simply the negative of the fractional part of the expression above. Since the vectors $(\bar{X}-X)$ and E_c satisfy the constraint conditions and since $\epsilon_b(t-j)$ is the same for all channels at the j-th lag,

$$K\left\{\sum_{i=1}^{M} \bar{x}(t-j) - x_i(t-j) + \epsilon_i(t-j) - \left[\frac{1}{M}\sum_{k=1}^{M}\epsilon_k(t-j)\right] + \epsilon_b(t-j)\right\} - KM\epsilon_b(t-j) = 0.$$

The fractional parts of the left-hand side of the equation must sum to an integer if the right-hand side is to equal zero:

$$-\left[KM\,\epsilon_b(t-j)\right]_F + \sum_{i=1}^{M} -\left[\delta_i(j)\right] = -I_j \quad ,$$

where I_j is an integer and where the subscript F denotes the fractional part of $KM\,\epsilon_b(t-j)$. This equation specifies a constraint on the roundoff errors $\delta_i(j)$. Any one of the components is a function of the other components and $KM\,\epsilon_b(t-j)$. Let it be the M-th component:

$$-\delta_M(j) = \left[KM\,\epsilon_b(t-j)\right]_F + \sum_{i=1}^{M-1}\delta_i(j) - I_j \quad .$$

$$= \left\{\left[KM\,\epsilon_b(t-j)\right]_F + \sum_{i=1}^{M-1}\delta_i(j)\right\}_F$$

$$+ \left\{\left[KM\,\epsilon_b(t-j)\right]_F + \sum_{i=1}^{M-1}\delta_i(j)\right\}_I - I_j \quad ,$$

where the subscript I denotes the integer part of the quantity inside the braces. Since the negative of the M-th component of the roundoff error lies between -1/2 and 1/2, it is equal to the fractional part of the right-hand side of the equation:

$$-\delta_M(j) = \left\{\left[KM\,\epsilon_b(t-j)\right]_F + \sum_{i=1}^{M-1}\delta_i(j)\right\}_F \quad .$$

Since $\delta_M(j)$ has no integer part, the integer part of the right-hand side must be zero, so that

$$I_j = \left\{ \left[KM\ \epsilon_b(t-j) \right]_F + \sum_{i=1}^{M-1} \delta_i(j) \right\}_I \quad .$$

Given probability distributions for $\epsilon_b(t-j)$ and the first $M-1$ components $\delta_i(j)$, the probability distribution of the integers I_j can be ascertained. The probability of any integer I_j is the probability that

$$I_j - 1/2 \le \left[KM\ \epsilon_b(t-j) \right]_F + \sum_{i=1}^{M-1} \delta_i(j) \le I_j + 1/2 \quad .$$

Another useful way of expressing I_j is

$$I_j = \left\{ \left[KM\, \epsilon_b(t-j) \right]_F + \left[KM\, \epsilon_b(t-j) \right]_I + \sum_{i=1}^{M-1} \delta_i(j) \right\}_I - \left[KM\, \epsilon_b(t-j) \right]_I$$

$$= \left\{ KM\, \epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j) \right\}_I - \left[KM\, \epsilon_b(t-j) \right]_I \quad .$$

Let L_j denote the integer part of $KM\, \epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j)$. Its probability is the probability that

$$L_j - 1/2 \le KM\, \epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j) \le L_j + 1/2 \quad .$$

An important random variable is the sum

$$\sum_{i=1}^{M} \delta_i(j) = I_j - \left[KM\,\epsilon_b(t-j)\right]_F$$

$$= L_j - KM\,\epsilon_b(t-j) ,$$

whose probability distribution depends on the probability of $KM\,\epsilon_b(t-j)$ and the conditional probability of L_j given $KM\,\epsilon_b(t-j)$.

Let z_i denote the filter-weight change

$$z_i = K\left\{\bar{x}(t-j) - x_i(t-j) + \epsilon_i(t-j) - \left[\frac{1}{M}\sum_{k=1}^{M} \epsilon_k(t-j)\right] + \epsilon_b(t-j)\right\}$$

of the i-th channel at lag j before roundoff. The roundoff error $\delta_i(j)$ depends only on the fractional part of z_i. Any integer change in z_i results in the same roundoff error $\delta_i(j) = -(z_i)_F$, so that the roundoff errors $\delta_i(j)$ for any point in any (M-1)-dimensional cell

$$m_i - 1/2 \le z_i \le m_i + 1/2 \qquad (i = 1, 2, \ldots, M-1;\ m_i \text{ an integer})$$

are the same as the roundoff errors $\delta_i(j)$ for the corresponding point in the cell

$$-1/2 \le z_i = (z_i)_F \le 1/2 \qquad (i = 1, 2, \ldots, M-1),$$

which includes the origin $(z_1, z_2, \ldots, z_{M-1}) = (0, 0, \ldots, 0)$. Consequently, the sum

$$\left[KM\,\epsilon_b(t-j)\right]_F + \sum_{i=1}^{M-1} \delta_i(j)$$

also remains invariant under integer translations of the variable z_i. Furthermore, the regions

$$I_j - 1/2 \leq \left[KM\epsilon_b(t-j)\right]_F + \sum_{i=1}^{M-1} \delta_i(j) \leq I_j + 1/2$$

corresponding to the integers I_j are identical in each of the (M-1)-dimensional cells. Figure III-23 illustrates this situation when $\left[KM\epsilon_b(t-j)\right]_F = 0$ for the three-channel case. $I_j = -1$ in the upper right corner of each cell, and $I_j = 1$ in the lower left corner of each cell. Any point within the shaded regions defined by $-1/2 \leq (z_1)_F + (z_2)_F \leq 1/2$ results in a value of zero for I_j. Non-zero values for $\left[KM\epsilon_b(t-j)\right]_F$ merely shift the boundaries for the integers I_j within each cell by altering the constant in the boundary equations $(z_1)_F + (z_2)_F =$ constant. Because of the constraints on the filter weights, the vectors $(z_1, z_2, \ldots, z_M)$ are concentrated in sheets whose z_M components differ by the scalar constant K. The central sheet corresponding to $KM\epsilon_b(t-j) = 0$ passes through the origin. When each of the components z_i is distributed over a sufficient number of filter-weight counts, the vectors $(z_1, z_2, \ldots, z_M)$ tend to become evenly distributed over the sheets within each of the (M-1)-dimensional cells

$$m_i - 1/2 \leq z_i \leq m_i + 1/2 \qquad (i = 1, 2, \ldots, M-1).$$

Accordingly, the vectors $(z_1, z_2, \ldots, z_M)$ projected onto the $(z_1, z_2, \ldots z_{M-1})$-plane also tend to become evenly distributed within each (M-1)-dimensional cell. Thus the conditional probability distribution of the vectors $\left[(z_1)_F, (z_2)_F, \ldots, (z_{M-1})_F\right]$ within each cell tends to become the same for all cells. As a result, the roundoff error components $\delta_i(j)$ $(i = 1, 2, \ldots, M-1)$ tend to be evenly distributed over the interval $[-1/2, 1/2]$. In the adaptive-filtering program used for this report, the changes z_i before rounding each

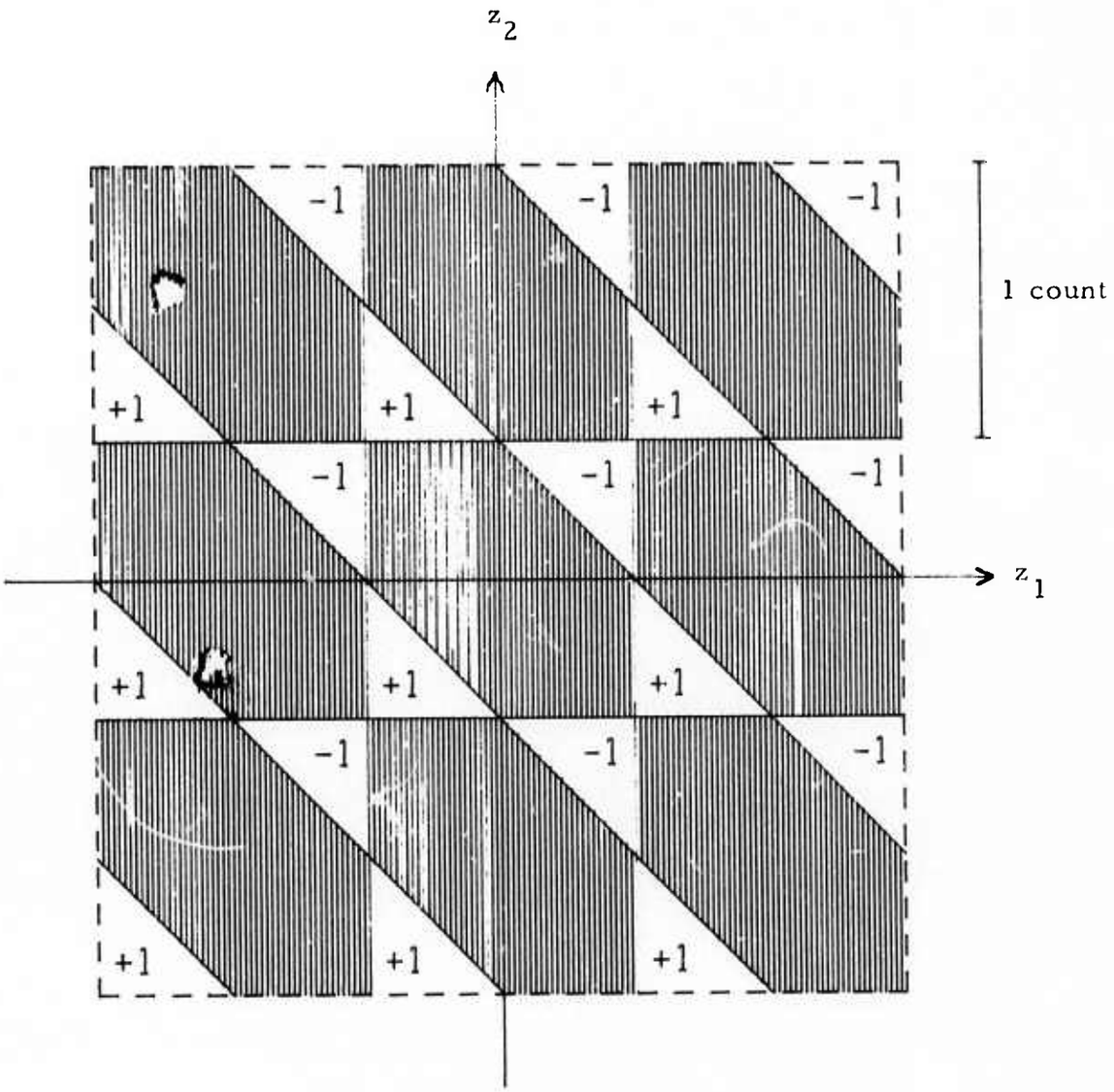


FIGURE III-23

REGIONS IN THE (z_I, z_2)-PLANE SATISFYING THE INEQUALITIES $I_j - 1/2 \leq \delta_1(j) + \delta_2(j) \leq I_j + 1/2$ (SHADED AREAS CORRESPOND TO THE INTEGER $I_j = 0$)

filter-weight component can assume values separated by only 2^{-32} filter-weight counts, so that the probability distribution for each of the first M-1 components $\delta_i(j)$ is adequately approximated by the probability density function

$$p\left[\delta_i(j)\right] = \begin{cases} 1 & |\delta_i(j)| < 1/2 \\ 0 & |\delta_i(j)| > 1/2 \end{cases}$$

whenever the filter-weight component changes are spread over a large number of filter-weight counts. The probability distribution for the beamsteer round-off error $\epsilon_b(t-j)$ was specified in Subsection B (Figure III-4). The resultant distributions for $KM\,\epsilon_b(t-j)$ and $[KM\,\epsilon_b(t-j)]_F$ are easily obtained from it once the scalar constant K is given. It is now possible to determine the probability distributions associated with the filter-weight roundoff vector D.

Let D_b denote the vector

$$D_b = \frac{1}{\sqrt{M}} \sum_{j=-N}^{N} \left[\sum_{i=1}^{M} \delta_i(j) \right] U_j \quad ,$$

where U_j is the unit vector perpendicular to the constraint plane in the subspace corresponding to the j-th lag (see Subsection C, page III-25). Since this vector is the sum of individual vectors perpendicular to the constraint space, it is likewise perpendicular to the constraint space. Let D_c be the vector $D_c = D - D_b$. It lies within the constraint space since, for each lag j ,

$$U_j^T D_c = U_j^T (D - D_b)$$

$$= \frac{1}{\sqrt{M}} \sum_{i=1}^{M} \delta_i(j) - \frac{1}{\sqrt{M}} \sum_{k=-N}^{N} \left[\sum_{i=1}^{M} \delta_i(j) \right] U_j^T U_k$$

$$= \frac{1}{\sqrt{M}} \left\{ \sum_{i=1}^{M} \delta_i(j) - \left[\sum_{i=1}^{M} \delta_i(j) \right] U_j^T U_j \right\}$$

$$= \frac{1}{\sqrt{M}} \left[\sum_{i=1}^{M} \delta_i(j) - \sum_{i=1}^{M} \delta_i(j) \right]$$

$$= 0 .$$

D therefore consists of a vector D_c within the constraint space and a vector D_b perpendicular to the constraint space.

The squared magnitude $|D_c|^2$ of the vector D_c is

$$D_c^T D_c = \sum_{j=-N}^{N} \sum_{i=1}^{M} \left\{ \delta_i(j) - \left[\frac{1}{M} \sum_{k=1}^{M} \delta_k(j) \right] \right\}^2$$

$$= \sum_{j=-N}^{N} \sum_{i=1}^{M} \left\{ \delta_i(j) - \frac{1}{M} \left[L_j - KM \epsilon_b(t-j) \right] \right\}^2$$

$$= \sum_{j=-N}^{N} \left\{ \sum_{i=1}^{M} \delta_i^2(j) - \frac{2}{M} \left[\sum_{i=1}^{M} \delta_i(j) \right] \left[L_j - KM\epsilon_b(t-j) \right] \right.$$

$$\left. + \frac{1}{M^2} \sum_{i=1}^{M} \left[L_j - KM\epsilon_b(t-j) \right]^2 \right\}$$

$$= \sum_{j=-N}^{N} \left\{ \sum_{i=1}^{M} \delta_i^2(j) - \frac{\left[L_j - KM\epsilon_b(t-j) \right]^2}{M} \right\} .$$

The random variable $\delta_M(j) = L_j - \left[KM\epsilon_b(t-j) + \sum_{k=1}^{M-1} \delta_k(j) \right]$ has the same probability distribution as the fractional part of the random variable $y = KM\epsilon_b(t-j) + \sum_{k=1}^{M-1} \delta_k(j)$. Since the random variables $\delta_k(j)$ $(k = 1, 2, \ldots, M-1)$ and $\epsilon_b(t-j)$ can be assumed mutually independent, the probability density function for the random variable y can be expressed as the convolution of the probability density function for one of the $M-1$ roundoff error components, say $\delta_\ell(j)$, with the probability density function for the random variable $x = KM\epsilon_b(t-j) + \sum_{\substack{k=1 \\ k \neq \ell}}^{M-1} \delta_k(j)$:

$$p(y) = \int_{y-1/2}^{y+1/2} p(x)\, dx .$$

Since

$$\sum_{n=-\infty}^{\infty} p(y+n) = \int_{-\infty}^{\infty} p(x)\, dx = 1 ,$$

the probability density function for the fractional part of the random variable $y = KM\epsilon_b(t-j) + \sum_{k=1}^{M-1} \delta_k(j)$ is a uniform density between -1/2 and 1/2. As a result, each of the M roundoff error components $\delta_i(j)$ (i= 1, 2, ..., M) has a uniform density between -1/2 and 1/2, even though the random variables $\delta_i(j)$ (i = 1, 2, ..., M) are not mutually independent. The probability distribution for each roundoff error component $\delta_i(j)$ is the same for all values of K. Such is not the case for the discrete random variable $L_j - KM\epsilon_b(t-j)$. A perspective drawing of its probability mass function as a function of K is shown in Figure III-24 for the six-channel case. Only the range K=0 to K=1/2 is shown since the probability mass function does not depend on the sign of K and repeats whenever K changes by an integer. Probability is concentrated in rays emanating from integer values L_j along the line K=0. Each ray corresponds to one of the seven integer values which $M\epsilon_b(t-j)$ can assume. Broken lines indicate rays where $M\epsilon_b(t-j) = \pm M/2$. The maximum probability mass of 0.092 is attained everywhere along the solid line directly above $L_j - KM\epsilon_b(t-j) = 0$. At points in the $(K, L_j - KM\epsilon_b(t-j))$-plane where rays intersect, the total probability of the value $L_j - KM\epsilon_b(t-j)$ is the sum of the individual probability mass functions associated with each of the intersecting rays. The spikes at points of intersection are omitted in order to avoid complicating Figure III-24 unduly. The probability of $L_j - KM\epsilon_b(t-j)$ given a particular value $KM\epsilon_b(t-j)$ is the probability that

$$L_j - KM\epsilon_b(t-j) - 1/2 \le \sum_{i=1}^{M-1} \delta_i(j) \le L_j - KM\epsilon_b(t-j) + 1/2 \ .$$

This probability is the integral of the probability density function for $\sum_{i=1}^{M-1} \delta_i(j)$ over a one-count-wide interval centered at $L_j - KM\epsilon_b(t-j)$. The integral is equal to the convolution of a uniform probability density between -1/2 and 1/2 with the probability density function for the random variable $\sum_{i=1}^{M-1} \delta_i(j)$, where the convolution integral is evaluated at $L_j - KM\epsilon_b(t-j)$. Since each of

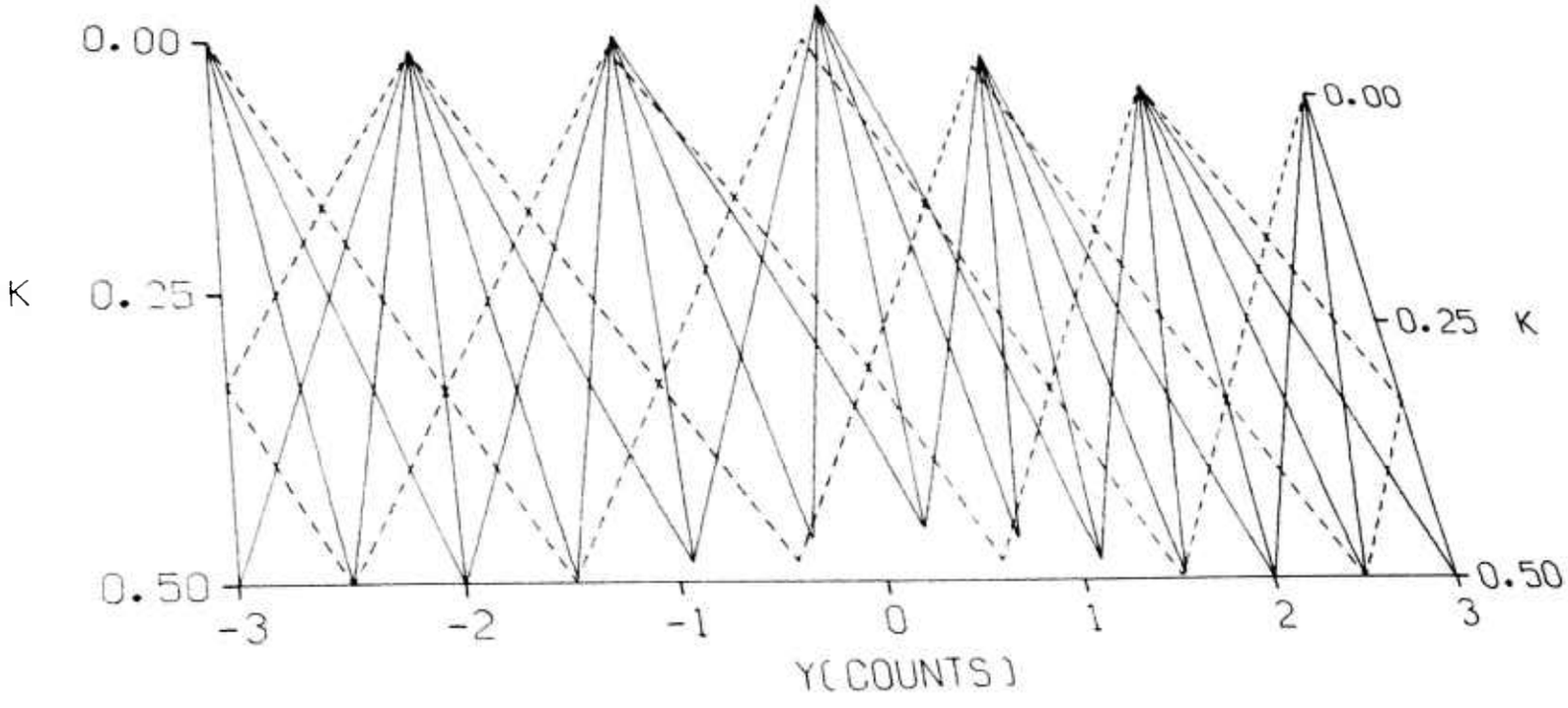


FIGURE III-24

PROBABILITY MASS FUNCTION FOR THE RANDOM VARIABLE
$Y = L_j - KM\epsilon_b(t-j)$ AS A FUNCTION OF K

the M-1 mutually independent random variables $\delta_i(j)$ is also assumed to have a uniform density between -1/2 and 1/2, the probability of $L_j - KM\epsilon_b(t-j)$ given $KM\epsilon_b(t-j)$ is the convolution of M uniform density functions (each equal to one over the interval -1/2 to 1/2), where the convolution is evaluated at $L_j - KM\epsilon_b(t-j)$. For this reason, the values $L_j - KM\epsilon_b(t-j)$ are confined to the interval (-M/2, M/2). To calculate the total probability of $L_j - KM\epsilon_b(t-j)$ at any point where it may occur, the M-fold convolution evaluated at $L_j - KM\epsilon_b(t-j)$ is multiplied by the sum of the probabilities of the values $M\epsilon_b(t-j)$ corresponding to the rays passing through the point in question. The probability of each value $M\epsilon_b(t-j)$ is 1/M unless $M\epsilon_b(t-j) = \pm M/2$, in which case it is 1/2M. The probability of $L_j - KM\epsilon_b(t-j)$ is zero whenever, at any particular value of K, no ray passes through the value $L_j - KM\epsilon_b(t-j)$. The probability of $L_j - KM\epsilon_b(t-j)$ given any value $M\epsilon_b(t-j)$ lies on an envelope equal to the M-fold convolution of unit-width, uniform density functions. Thus the broken lines in Figure III-24 are located on an envelope half as high as the envelope containing the solid lines.

To determine the mean of the squared magnitude $|D_c|^2$ of the vector D_c, the expression for $|D_c|^2$ can be rewritten:

$$D_c^T D_c = \sum_{j=-N}^{N} \left\{ \left[\sum_{i=1}^{M} \delta_i^2(j) \right] - \frac{1}{M} \left[\sum_{i=1}^{M} \sum_{k=1}^{M} \delta_i(j)\, \delta_k(j) \right] \right\}$$

$$= \sum_{j=-N}^{N} \left\{ \frac{M-1}{M} \left[\sum_{i=1}^{M} \delta_i^2(j) \right] - \frac{1}{M} \left[\sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \neq i}}^{M} \delta_i(j)\, \delta_k(j) \right] \right\}$$

$$= \frac{1}{M} \sum_{j=-N}^{N} \sum_{i=1}^{M-1} \sum_{k=i+1}^{M} \left[\delta_i^2(j) + \delta_k^2(j) - 2\,\delta_i(j)\, \delta_k(j) \right]$$

$$= \frac{1}{M} \sum_{j=-N}^{N} \sum_{i=1}^{M-1} \sum_{k=i+1}^{M} \left[\delta_i(j) - \delta_k(j) \right]^2 \quad .$$

The mean of $|D_c|^2$ is $1/M$ times the sum of the means for the individual squared terms $\left[\delta_i(j) - \delta_k(j)\right]^2$. This fact does not depend on the mutual independence of the squared terms. Since $\delta_i(j)$ and $-\delta_k(j)$ have zero means, the mean of $\left[\delta_i(j) - \delta_k(j)\right]^2$ is the variance of the random variable $\delta_i(j) - \delta_k(j)$. Any subset of 2 to M-1 roundoff error components at the j-th lag contains mutually independent elements, so that

$$\text{Var} \left[\delta_i(j) - \delta_k(j) \right] = \text{Var} \left[\delta_i(j) \right] + \text{Var} \left[\delta_k(j) \right]$$

whenever $M \geq 3$. For a uniform probability density between -1/2 and 1/2, the variance is 1/12, so that the mean of $|D_c|^2$ is

$$\overline{D_c^T D_c} = \frac{2N+1}{M} \cdot \frac{(M-1)M}{2} \cdot \frac{2}{12} = \frac{(M-1)(2N+1)}{12}$$

whenever $M \geq 3$. Unless M=2, the mean of $|D_c|^2$ and the contribution to the mean from the subspace corresponding to the j-th lag do not depend on K. A simple corollary results from this fact. Since

$$\frac{1}{M} \overline{\left[L_j - KM \epsilon_b(t-j) \right]^2} = \left[\sum_{i=1}^{M} \overline{\delta_i^2(j)} \right] - \frac{(M-1)}{12} = \frac{1}{12} \quad ,$$

the variance of the random variable $L_j - KM \epsilon_b(t-j)$ is M/12 (independent of K) whenever $M \geq 3$.

The squared magnitude $|D_b|^2$ of the vector D_b is

$$D_b^T D_b = \frac{1}{M} \sum_{j=-N}^{N} \sum_{k=-N}^{N} \left[\sum_{i=1}^{M} \delta_i(j) \right] \left[\sum_{i=1}^{M} \delta_i(k) \right] U_j^T U_k$$

$$= \frac{1}{M} \sum_{j=-N}^{N} \left[\sum_{i=1}^{M} \delta_i(j) \right]^2$$

$$= \frac{1}{M} \sum_{j=-N}^{N} \left[L_j - KM \epsilon_b(t-j) \right]^2$$

Its mean is equal to $(2N+1)/12$ whenever $M \geq 3$. The mean of the squared magnitude $|D|^2 = |D_c|^2 + |D_b|^2$ is $M(2N+1)/12$ for the same range of M values. The vector D_b, however, combines with the vector KE_b in such a way that

$$|D_b + KE_b|^2 = \frac{1}{M} \sum_{j=-N}^{N} \left[L_j - KM \epsilon_b(t-j) + KM \epsilon_b(t-j) \right]^2$$

$$= \frac{1}{M} \sum_{j=-N}^{N} L_j^2 \ ,$$

where the probability of any integer L_j is the probability that

$$L_j - 1/2 \leq KM \epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j) \leq L_j + 1/2 \ .$$

It depends solely on the probability distribution for the random variable $KM \epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j)$, whose probability density function is pictured in

Figure III-25 for the values K=0 to K=6. By the time K rises to the value 2.5, the fingers corresponding to each of the seven possible values of $\epsilon_b(t-j)$ begin to separate. Once K is 5.0 or above, they are completely distinct. The maximum probability density of 0.599 is attained at K=0, $KM\,\epsilon_b(t-j) + \sum_{j=1}^{M-1} \delta_i(j) = 0$. For values of K greater than 2.5, the maximum probability density is 0.100 on the inner five fingers, 0.050 on the outer two. The probability mass function for L_j also varies with K. Since the probability of L_j is the probability that the random variable $KM\,\epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j)$ lies within 1/2 count of L_j, the probability contained within the one-count-wide lanes straddling integers in Figure III-25 condenses to form the probability mass function for L_j, which is plotted as a function of K in Figure III-26. The maximum probability mass of 0.550 occurs at the origin. For values of K greater than 3.0, the maximum probability mass is 0.092 on the inner five fingers, 0.046 on the outer two. Note that the finger corresponding to $\epsilon_b(t-j) = 0$ in Figure III-25 necessarily results in a non-zero probability for the integers $L_j = -1$, $L_j = 0$, and $L_j = 1$ at all values of K in Figure III-26. This fact is important in part 2 of this subsection. The random variable L_j is of interest in its own right. It represents the amount by which the error $\epsilon_b(t-j)$ and the errors $\delta_i(j)$ at the j-th lag shift the sum $\sum_{i=1}^{M} a_i(j)$ either toward or away from the j-th lag constraint condition. At the j-th lag, it completely specifies the probability distribution of the random walk away from the constraint condition.

The mean of $|D_b + KE_b|^2$ is

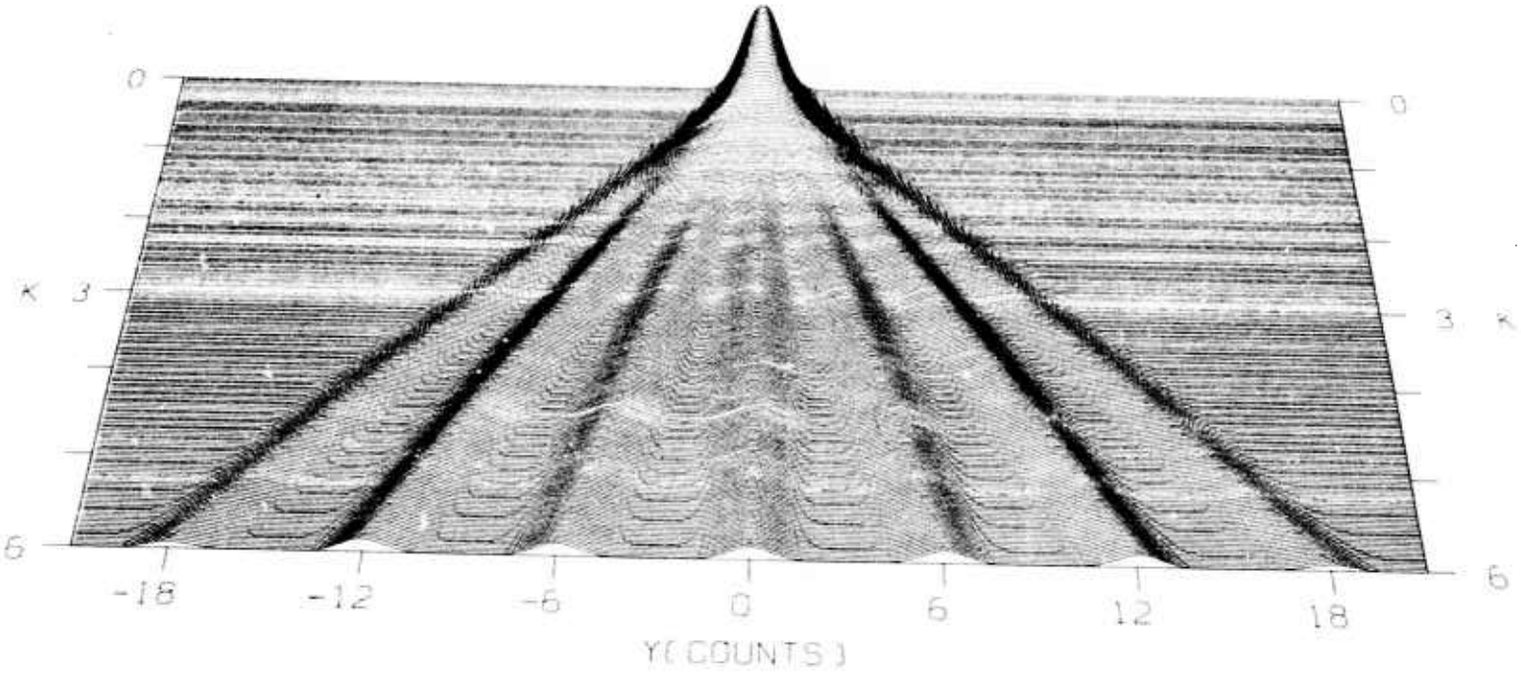


FIGURE III-25

PROBABILITY DENSITY FUNCTION FOR THE RANDOM VARIABLE
$Y = KM\,\epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j)$ AS A FUNCTION OF K

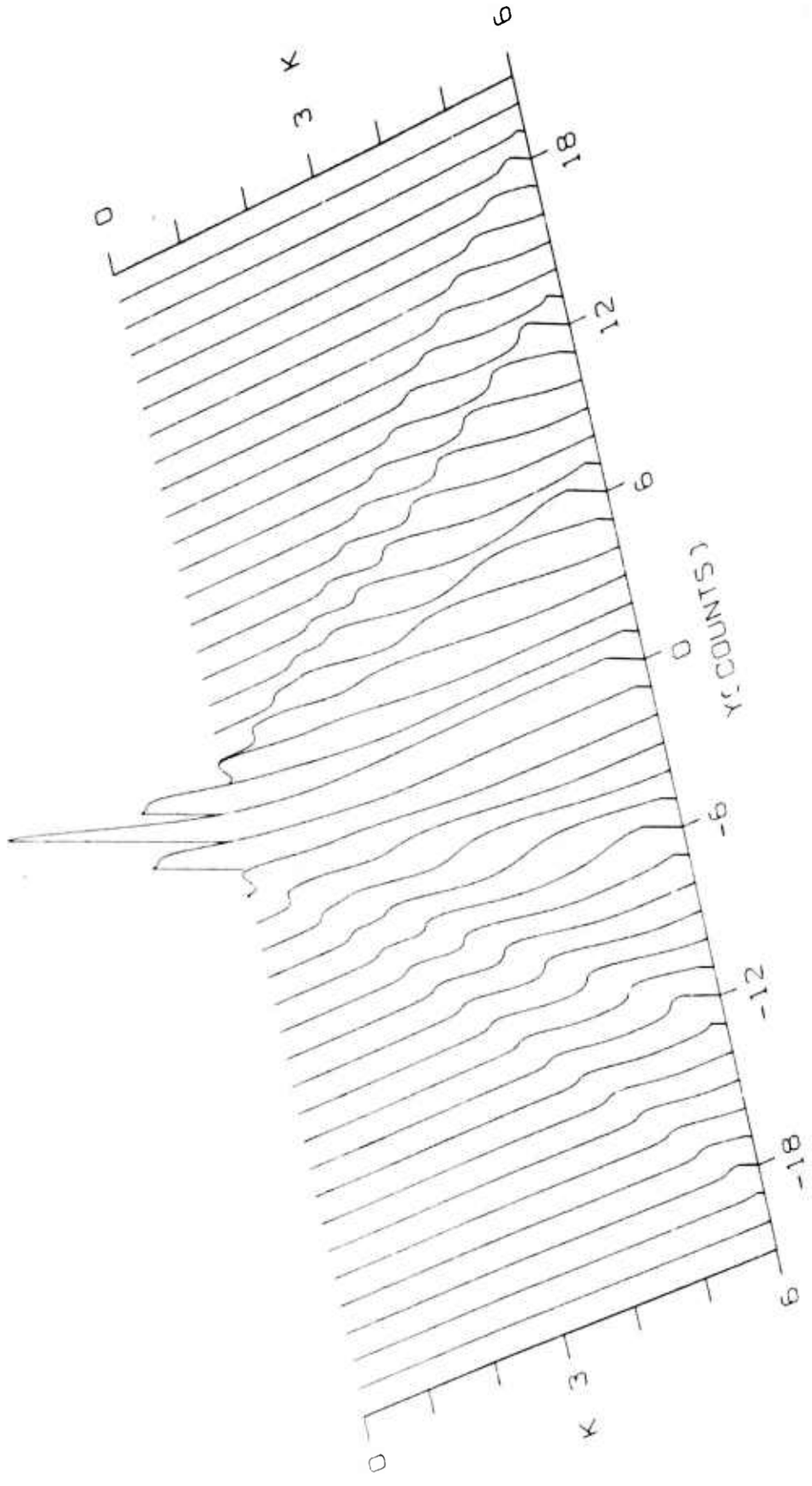


FIGURE III-26

PROBABILITY MASS FUNCTION FOR THE RANDOM VARIABLE $Y=L_j$ AS A FUNCTION OF K

$$\overline{|D_b + KE_b|^2} = \frac{1}{M} \sum_{j=-N}^{N} \left\{ \overline{\left[L_j - KM\epsilon_b(t-j)\right]^2} \right.$$

$$\left. + 2\overline{\left[L_j - KM\epsilon_b(t-j)\right] KM\epsilon_b(t-j)} + \overline{\left[KM\epsilon_b(t-j)\right]^2} \right\} .$$

Since M-1 of the roundoff error components $\delta_i(j)$ at the j-th lag are independent of $KM\epsilon_b(t-j)$,

$$2\overline{\left[L_j - KM\epsilon_b(t-j)\right] KM\epsilon_b(t-j)}$$

$$= 2\overline{\left[\sum_{i=1}^{M} \delta_i(j)\right] KM\epsilon_b(t-j)}$$

$$= 2\overline{\delta_M(j)\, KM\epsilon_b(t-j)}$$

$$= -2\overline{\left[KM\epsilon_b(t-j) + \sum_{i=1}^{M} \delta_i(j)\right]_F KM\epsilon_b(t-j)} .$$

Whenever $M \geq 2$, the probability distribution for $\delta_M(j)$ is uniform between -1/2 and 1/2 and does not depend on the value of $KM\epsilon_b(t-j)$, so that

$$2\overline{\left[L_j - KM\epsilon_b(t-j)\right] KM\epsilon_b(t-j)} = 2\,\overline{\delta_M(j)}\;\overline{KM\epsilon_b(t-j)} = 0$$

and

$$\overline{|D_b + KE_b|^2} = \frac{1}{M} \sum_{j=-N}^{N} \left\{ \overline{\left[L_j - KM\epsilon_b(t-j)\right]^2} + \overline{\left[KM\epsilon_b(t-j)\right]^2} \right\} .$$

Whenever $M \geq 3$, the variance of the random variable $L_j - KM\epsilon_b(t-j)$ is $M/12$ and does not depend on K, so that

$$\overline{|D_b + KE_b|^2} = \frac{2N+1}{12} + K^2 \sum_{j=-N}^{N} M\,\overline{\epsilon_b^2(t-j)}$$

$$= \frac{2N+1}{12} + K^2\,\overline{|E_b|^2} \quad .$$

For the case of 6 channels and 31 filter points, the mean of $|E_b|^2$ is 16.361, and

$$\overline{|D_b + KE_b|^2} = 2.583 + 16.361\,K^2 \quad .$$

In order to derive these results, it was presumed that the filter-weight change components z_i were distributed over a sufficient number of filter-weight counts that the vectors $(z_1, z_2, \ldots, z_M)$ tended to become evenly distributed over sheets within each of the $(M-1)$-dimensional cells

$$m_i - 1/2 \leq z_i \leq m_i + 1/2 \qquad (i=1, 2, \ldots, M-1)$$

in the subspace corresponding to any one of the $2N+1$ lags of the filter set. As a result of this premise, the roundoff error components $\delta_i(j)$ assumed probability distributions independent of both the prefiltered vertical-component data values $x_i(t-j)$ and errors $\epsilon_i(t-j)$. The lack of correlation between the prefiltered vertical-component errors $\epsilon_i(t-j)$ and filter-weight roundoff errors $\delta_i(j)$ implies that the mean of $|D_c + KE_c|^2$ is

$$\overline{|D_c + KE_c|^2} = \overline{|D_c|^2} + K^2\,\overline{|E_c|^2}$$

$$= \frac{(M-1)(2N+1)}{12} + K^2\,\overline{|E_c|^2} \quad .$$

This mean is $12.917 + 28.071\ K^2$ for unscaled data, $12.917 + 1081.952\ K^2$ for data scaled by 16. The mean of the squared magnitude of the error vector $D + KE$ is

$$\overline{|D_b + D_c + K(E_b + E_c)|^2} = \overline{|D_b + KE_b|^2} + \overline{|D_c + KE_c|^2}$$

$$= \overline{|D_b|^2} + \overline{|D_c|^2} + K^2\left(\overline{|E_b|^2} + \overline{|E_c|^2}\right)$$

$$= \frac{M(2N+1)}{12} + K^2\left(\overline{|E_b|^2} + \overline{|E_c|^2}\right)$$

since both D_b and E_b are perpendicular to either D_c or E_c. The resultant mean is $15.5 + 44.432\ K^2$ for unscaled data, $15.5 + 1098.313\ K^2$ for data scaled by 16.

In the same way that an estimate was obtained in Subsection C for the mean angle between the vector $(\overline{X}-X)$ and the vector $(\overline{X}-X) + E_c$, an estimate can be made for the mean $\overline{\theta}$ of the angle θ between the vectors $(\overline{X}-X)$ and $\text{sgn}[y(t)]\ \text{sgn}[y'(t)]\ (\overline{X}-X + E_c + D_c/K)$:

$$\overline{\theta} \approx \tan^{-1}\sqrt{\frac{\overline{|D_c + KE_c|^2}}{K^2 E[|\overline{X}-X|^2]}}$$

$$= \tan^{-1}\sqrt{\frac{\overline{|D_c|^2}}{K^2 E[|\overline{X}-X|^2]} + \frac{\overline{|E_c|^2}}{E[|\overline{X}-X|^2]}}$$

$$\approx \tan^{-1}\sqrt{\frac{q^2 E[|\overline{X}-X + E_b + E_c|^4]\ \overline{|D_c|^2}}{4K_s^2\ \overline{[y'(t)]^2}\ E[|\overline{X}-X|^2]} + \frac{\overline{|E_c|^2}}{E[|\overline{X}-X|^2]}}$$

$$\approx \tan^{-1} \sqrt{\frac{q^2 \overline{|D_c|^2}\ E[|\bar{X}-X|^2]}{4K_s^2 \overline{[y'(t)]^2}} + \frac{\overline{|E_c|^2}}{E[|\bar{X}-X|^2]}} \quad .$$

In this estimate, the angle between $D_c + KE_c$ and $(\bar{X}-X)$ has been assumed to be 90°. In addition, the difference between $E[|\bar{X}-X+E_b+E_c|^2]$ and $E[|\bar{X}-X|^2]$ has been assumed negligible. Furthermore, the probability distributions for $|D_c|^2$, $|\bar{X}-X|^2$, $[y'(t)]^2$, and $|E_c|^2$ are presumed to be sharply peaked around their respective mean values. In view of all these approximations, the estimate here is not likely to be as accurate as the corresponding estimate in Subsection C. For the four-hour noise sample from day 238 of 1970 which has been used throughout this section, the mean of $|\bar{X}-X|^2$ was 9693 counts^2 for unscaled data, 2,481,422 counts^2 for data scaled by 16. When substitutions are made for values which do not depend on the convergence factor K_s, the estimated mean angle becomes

$$\bar{\theta} \approx \tan^{-1} \sqrt{\frac{(2^{-32})(12.917)(9693)}{4K_s^2 \overline{[y'(t)]^2}} + \frac{\overline{|E_c|^2}}{(9693)}}$$

$$= \tan^{-1} \sqrt{\frac{7.2878 \times 10^{-6}}{K_s^2 \overline{[y'(t)]^2}} + 2.8904 \times 10^{-3}}$$

for unscaled data,

$$\bar{\theta} = \tan^{-1} \sqrt{\frac{7.2878 \times 10^{-6}}{K_s^2 \overline{[y'(t)]^2}} + 4.3602 \times 10^{-4}}$$

for data scaled by 16. Here the mean adaptive filter output power $[y'(t)]^2$

for data scaled by 16 is expressed in terms of the equivalent value corresponding to unscaled data. Table III-8 gives the adaptive filter output power for data scaled by 16 (after division by 256) as a function of the convergence factor K_s. The resulting estimated mean of the angle θ for unscaled data and data scaled by 16 appears in the two rightmost columns as a function of the convergence factor K_s. Figure III-27 graphs the estimated mean angles $\overline{\theta}$ of Table III-8 as a function of the convergence factor K_s for unscaled data and data scaled by 16. When the adaptive filter output $y(t)$ without errors and the adaptive filter output $y'(t)$ with errors are non-zero and have the same sign, θ is the angle between the vector $(A^{new}-A^{old})$ without any digitization or roundoff error and the projection onto the constraint space of the vector $(A^{new}-A^{old})$ with digitization and roundoff error. The mean adaptive filter output power $[y'(t)]^2$ was taken from computer runs where the data was scaled by 16. At higher convergence rates, the value $[y'(t)]^2$ for unscaled data is likely to be larger than for scaled data. However, the term $|E_c|^2 / E[|\overline{X}-X|^2]$ predominates at high values of K_s. At lower convergence rates, the adaptive filter output power is likely to be approximately the same for unscaled data and data scaled by 16. Since no runs were made with unscaled data, the estimate of $\overline{\theta}$ in the top curve of Figure III-27 is the best possible without $[y'(t)]^2$ measurements for unscaled data. The estimates of $\overline{\theta}$ in Figure III-27 are for the angle θ in the constraint space before any attempt to reduce discrepancies between the filter weights $a_i(j)$ and the constraint conditions at each of the $2N+1$ lags in the adaptive filter set. A discussion of the angle of error after corrective procedures designed to force reasonable agreement with the constraint conditions must be postponed until part 2 of this subsection.

The effect on the mean angle of error $\overline{\theta}$ due to a change in quantization level may be evaluated in terms of the approximation for $\overline{\theta}$. If the quantization level is halved by adding one additional bit to the filter-weight

TABLE III-8

ESTIMATED MEAN $\bar{\theta}$ OF THE ANGLE BETWEEN $(\bar{X}\text{-}X)$ AND $\text{sgn}\,[y(t)]\ \text{sgn}\,[y'(t)]\ (\bar{\bar{X}}\text{-}X + E_c + D_c/K)$ AS A FUNCTION OF THE CONVERGENCE FACTOR K_s (USING NOISE DATA FROM DAY 238 OF 1970)

Convergence Factor K_s	Adaptive Filter Power (counts^2)	Estimated Mean of the Angle θ	
		Unscaled Data	Data Scaled by 16
0.002	11.0875	22° 15'	22° 6'
0.003	10.4431	15° 50'	15° 36'
0.004	10.0444	12° 23'	12° 5'
0.005	9.7905	10° 15'	9° 52'
0.007	9.4075	7° 47'	7° 16'
0.010	9.0444	5° 58'	5° 16'
0.015	8.6332	4° 40'	3° 42'
0.02	8.3197	4° 4'	2° 55'
0.03	7.7843	3° 35'	2° 12'
0.05	6.9377	3° 18'	1° 41'
0.07	6.2887	3° 12'	1° 29'
0.10	5.6118	3° 9'	1° 22'
0.15	5.0118	3° 7'	1° 17'
0.20	4.7441	3° 6'	1° 15'
0.25	4.6290	3° 5'	1° 14'
0.30	4.5788	3° 5'	1° 13'
0.35	4.5733	3° 5'	1° 13'
0.40	4.5799	3° 5'	1° 13'
0.50	4.6254	3° 5'	1° 12'

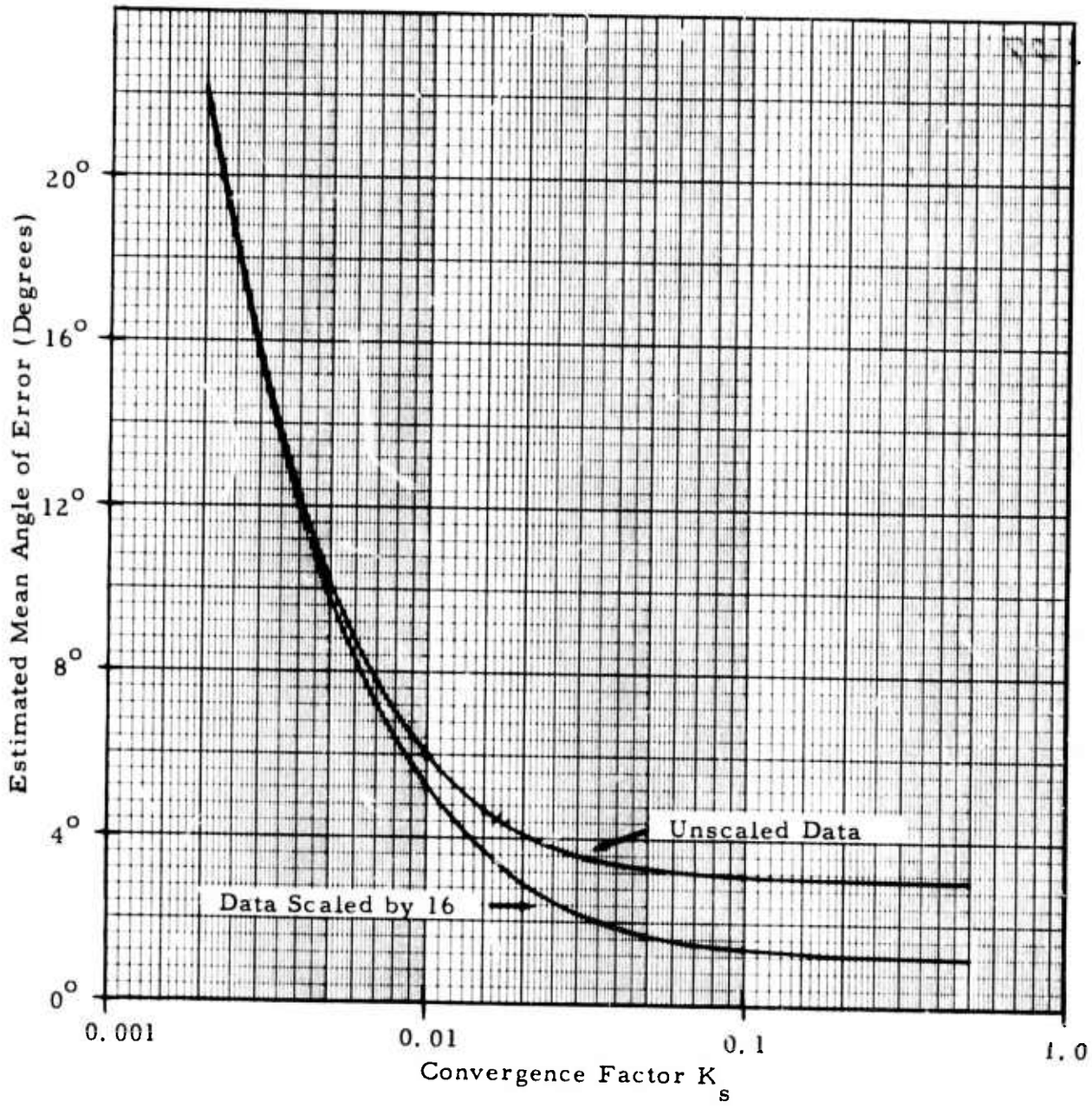


FIGURE III-27

ESTIMATED MEAN $\bar{\theta}$ OF THE ANGLE BETWEEN $(\bar{X}-X)$ AND $\text{sgn}\,[y(t)]\,\text{sgn}\,[y'(t)]\,(\bar{X}-X+E_c+D_c/K)$ AS A FUNCTION OF CONVERGENCE RATE (USING NOISE DATA FROM DAY 238 OF 1970)

representation, then the same mean angle $\bar{\theta}$ is achieved at about half the original convergence factor K_s if the change in the adaptive-filter output power can be ignored. For the data in Table III-8, the change in the adaptive-filter output power is near 10% when K_s changes by a factor of 2, so that the gross effect of scaling the quantization level is to multiply the abscissa K_s in Figure III-27 by an equivalent amount.

2. Error in the Constraint Conditions

If a signal $s(t)$ from the look direction propagates across the array from which the data used in an adaptive-filtering process originates, the output of the adaptive filter is

$$\sum_{j=-N}^{N} \left[\sum_{i=1}^{M} a_i(j) \right] s(t-j) \quad .$$

In effect, a filter with weights $\sum_{i=1}^{M} a_i(j)$ $(j=-N, \ldots, -1, 0, 1, \ldots, N)$ is applied to the signal. The constraint conditions

$$\sum_{i=1}^{M} a_i(j) = a(j)$$

determine the frequency response of the filter set. In the adaptive-filtering process implemented for this report, a white frequency response is specified by setting $a(j) = \delta_{jo}$ (where δ_{jo} is the Kronecker delta operator). The effect of not satisfying the constraint conditions is to alter the frequency response of the adaptive processor to a signal from the look direction.

After a large number of iterations, a random walk away from the constraint space takes place in the absence of an error-correcting procedure. In

the case of floating-point data, many iterations may elapse before the damage is serious. Ultimately, however, the filter vector A must be pulled back toward the constraint space. The error correction may be performed periodically, or it may be executed at each iteration with a technique similar to O. L. Frost's (Frost, 1972). With integer arithmetic, the roundoff errors are normally at least an order of magnitude greater than for floating-point arithmetic, so that frequent corrective action is necessary.

At each iteration, the probability distribution of the change in the sum $\sum_{i=1}^{M} a_i(j)$ for the j-th lag is a discrete integer distribution whose probability mass function was depicted in Figure III-26. The probability of each possible integer shift L_j is the probability that

$$L_j - 1/2 \leq KM\epsilon_b(t-j) + \sum_{i=1}^{M-1} \delta_i(j) \leq L_j + 1/2 \quad .$$

On the first iteration, the squared distance from the constraint space (in squared filter-weight counts) becomes

$$|D_b + KE_b|^2 = \frac{1}{M} \sum_{j=-N}^{N} L_j^2 \quad ,$$

whose contribution from the j-th lag is L_j^2/M. In order to bring the filter vector back to the constraint space along a vector perpendicular to the constraint space, the quantity L_j/M must be subtracted from each component of the filter vector. This procedure is possible only at those lags where L_j is a multiple of M, the number of channels. By subtracting the appropriate integer, however, it is possible to reduce the discrepancy between the constraint condition and the filter weights at the j-th lag to the point where

$$-(M-1) \leq \delta_{jo} - \sum_{i=1}^{M} a_i(j) \leq M-1 \ .$$

Let J denote the discrepancy. In order to satisfy the constraint conditions with minimum movement inside the constraint plane, $\mathrm{sgn}(J)$ must be subtracted from J of the M components at one lag, while $M-|J|$ components must be left alone. At any one lag, the squared magnitude of the resultant displacement within the constraint plane is

$$|J| \left(\frac{M - |J|}{M}\right)^2 + \frac{(M - |J|)J^2}{M^2}$$

$$= \frac{|J| \ (M - |J|)}{M}$$

For the integers such that $1 \leq |J| < M/2$, the squared magnitude is greater than the previous squared magnitude J^2/M for the vector perpendicular to the constraint plane. It is the same when $|J| = M/2$, and less when $M/2 < |J| < M$. This procedure was employed by the adaptive-filtering system which provides the results for this report. It converts error perpendicular to the constraint space into error within the constraint space. This feature is an annoying disadvantage of the procedure.

If C is the error vector due to conversion of error perpendicular to the constraint space into error within the constraint space, the mean of $|C|^2$ is

$$\overline{|C|^2} = \left(\frac{2N+1}{M}\right) \overline{|J| \ (M - |J|)} \ .$$

If $|J|$ were always equal to $M/2$, $\overline{|C|^2}$ would reach the maximum value

$$|C|^2 = M(2N+1)/4 \quad .$$

It is difficult to envision a situation where the mean value would be equal to the maximum. However, if the integer shifts L_j were distributed in such a way that the values $J = L_j$ (modulo M) were equally likely at the values $J = 0, 1, \ldots, M-1$, the mean of $|C|^2$ would be

$$\overline{|C|^2} = \frac{(2N+1)}{M^2} \sum_{J=0}^{M-1} |J| \, (M - |J|) = \frac{(2N+1)(M^2 - 1)}{6M} \quad .$$

Such a distribution for the integer shifts L_j would occur if K were sufficiently high. The least possible value for $\overline{|C|^2}$ would occur at K=0, where $L_j = \pm 1$ about half the time and $L_j = 0$ about half the time (when M=6). At K=0,

$$\overline{|C|^2} \approx \frac{(2N+1)(1)(5)}{2\,M} = \frac{31 \cdot 5}{12} = 12.917 \quad .$$

The mean angle $\bar{\theta}$ between the vector $(\bar{X}-X)$ and the vector $\text{sgn}\,[y(t)]\,\text{sgn}\,[y'(t)]\,(\bar{X}-X + E_c + D_c/K + C/K)$ can be estimated as before:

$$\bar{\theta} \approx \tan^{-1} \sqrt{\frac{q^2(\overline{|D_c|^2} + \overline{|C|^2})\, E[|\bar{X}-X|^2]}{4K_s^2\, \overline{[y'(t)]^2}} + \frac{\overline{|E_c|^2}}{E[|\bar{X}-X|^2]}} \quad .$$

At K=0, $\overline{|D_c|^2} + \overline{|C|^2} \approx 25.833$. When the values $J = L_j$ (modulo M) are equally likely at 0, 1, ..., M-1, $\overline{|D_c|^2} + \overline{|C|^2} = 43.056$. If $|J|$ were always M/2, $\overline{|D_c|^2} + \overline{|C|^2}$ would be 59.417. The value 43.056 will be used. When $E[|\bar{X}-X|^2] = 9693 \text{ counts}^2$,

$$\bar{\theta} \approx \tan^{-1} \sqrt{\frac{2.4292 \times 10^{-5}}{K_s^2 \overline{[y'(t)]^2}} + 2.8904 \times 10^{-3}}$$

for unscaled data, and

$$\bar{\theta} \approx \tan^{-1} \sqrt{\frac{2.4292 \times 10^{-5}}{K_s^2 \overline{[y'(t)]^2}} + 4.3602 \times 10^{-4}}$$

for data scaled by 16. Table III-9 lists these mean angles $\bar{\theta}$ for unscaled data and data scaled by 16 as a function of the convergence factor K_s. Figure III-28 is a graph of the values in Table III-9. If $y'(t)$ is a non-zero value with the same sign as $y(t)$, θ is the angle of error in $(A^{new} - A^{old})$ after the constraint-condition correction procedure which was actually implemented. After the correction procedure is executed, the constraint conditions are satisfied exactly and all error lies within the constraint space. At the convergence factor $K_s = 0.005$ utilized most frequently in processing data for this report, the mean angle of error is less than 18^{o}. This angle is tolerable, although more than desired. The result of the correction procedure actually employed is almost to double the angle of error within the constraint space. In effect, one bit in the filter-weight representation is lost when the mean angles of Figure III-28 are compared with those of Figure III-27.

In view of the very small error angles at high convergence rates, the drop in noise reduction from 6 dB to 2 dB at $K_s = 0.30$ for one summer noise sample when the data scale factor was reduced from 16 to I is almost certainly due to zero $y'(t)$ values and sign differences between $y'(t)$ and $y(t)$.

An alternative procedure which does not add to the error within the constraint space is to round the value

TABLE III-9

ESTIMATED MEAN $\bar{\theta}$ OF THE ANGLE BETWEEN $(\bar{X}\text{-}X)$ AND $\text{sgn}\,[y(t)]\ \text{sgn}\,[y'(t)]\ (\bar{X}\text{-}X + E_c + D_c/K + C/K)$ AS A FUNCTION OF THE CONVERGENCE FACTOR K_s (USING NOISE DATA FROM DAY 238 OF 1970)

Convergence Factor K_s	Adaptive Filter Power (counts^2)	Estimated Mean of the Angle θ	
		Unscaled Data	Data Scaled by 16
0.002	11.0875	36° 35'	36° 31'
0.003	10.4431	27° 5'	26° 58'
0.004	10.0444	21° 26'	21° 16'
0.005	9.7905	17° 43'	17° 31'
0.007	9.4075	13° 16'	12° 59'
0.010	9.0444	9° 47'	9° 23'
0.015	8.6332	7° 4'	6° 29'
0.02	8.3197	5° 45'	5° 0'
0.03	7.7843	4° 34'	3° 35'
0.05	6.9377	3° 45'	2° 27'
0.07	6.2887	3° 28'	2° 0'
0.10	5.6118	3° 18'	1° 41'
0.15	5.0118	3° 11'	1° 28'
0.20	4.7441	3° 9'	1° 22'
0.25	4.6290	3° 7'	1° 18'
0.30	4.5788	3° 7'	1° 16'
0.35	4.5733	3° 6'	1° 15'
0.40	4.5799	3° 6'	1° 14'
0.50	4.6254	3° 5'	1° 13'

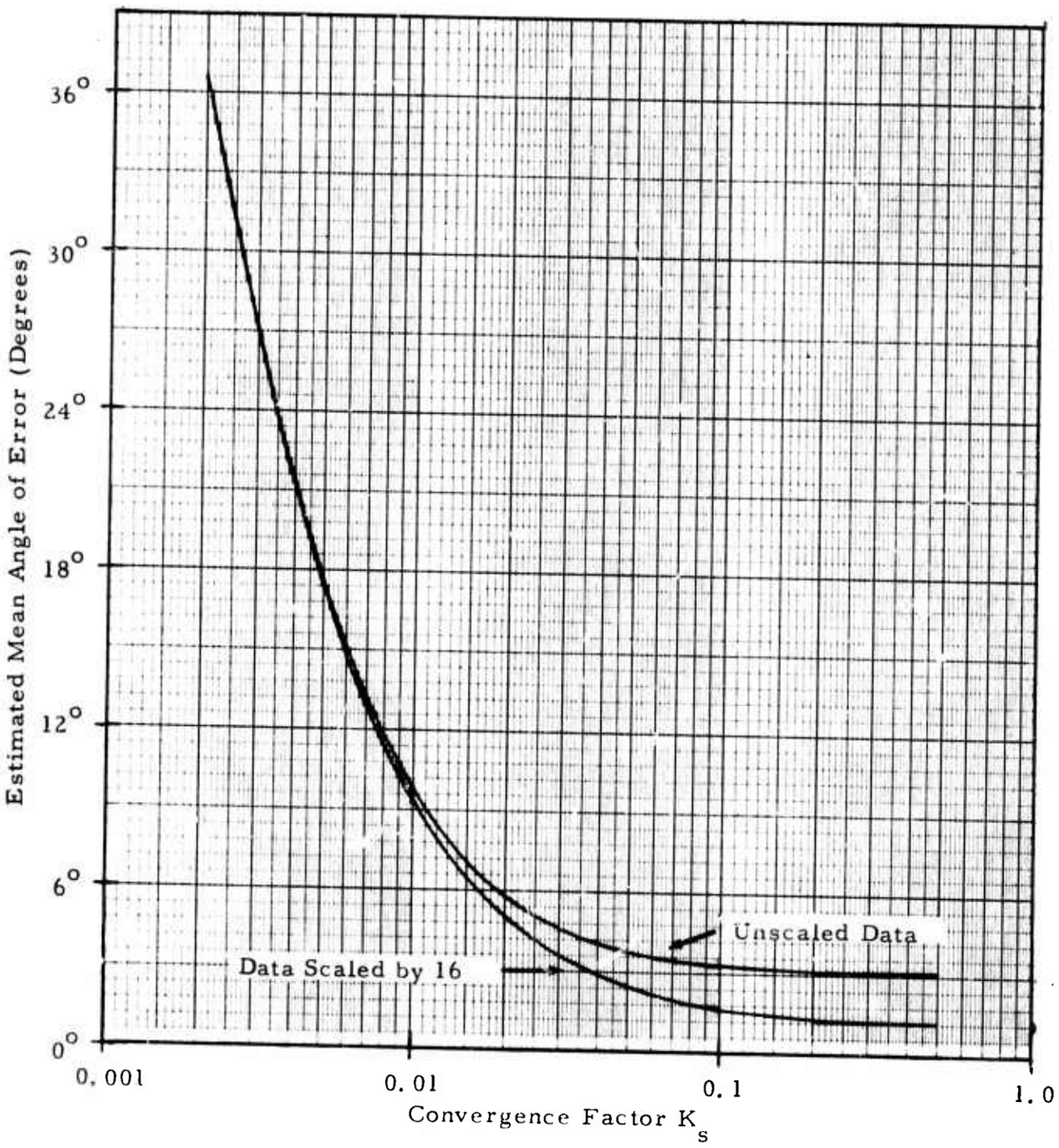


FIGURE III-28

ESTIMATED MEAN $\bar{\theta}$ OF THE ANGLE BETWEEN $(\bar{X}\text{-}X)$ AND $\text{sgn}\,[y(t)]\,\text{sgn}\,[y'(t)](\bar{X}\text{-}X + E_c + D_c/K + C/K)$ AS A FUNCTION OF CONVERGENCE RATE (USING NOISE DATA FROM DAY 238 OF 1970)

$$\left\{\left[\sum_{i=1}^{M} a_i(j)\right] - \delta_{jo}\right\} / M$$

to the nearest integer and subtract the resulting integer from each component at the j-th lag. When M is an even integer and the discrepancy is equal to M/2 (modulo M), the value to be subtracted is rounded to the integer with the smallest absolute value to avoid unproductive computations. If M were 6 and the discrepancy were 4, this error-correction procedure would subtract 1 from each component and produce a discrepancy of -2. The squared distance to the constraint plane would fall from 16/6 to 4/6. The other procedure would add 8/6 to the squared error within the constraint plane. In fact, the suggested procedure always results in less error than the one actually utilized unless the discrepancy is congruent (modulo M) with zero or M/2. The suggested procedure limits the deviation from the constraint conditions to the set of integers with absolute value less than (M+1)/2. Such a small error should only minimally affect the frequency response of the adaptive filter to a signal from the look direction when a 16-bit representation of the filter weights is employed.

If the suggested procedure were implemented, the probability of each possible integer discrepancy between the filter weights and the constraint condition at any lag could be described in terms of a Markov chain. Figure III-29 diagrams the set of possible integers and transitions between integers for the five-channel case. It is representative of the situation for an odd number of channels. Figure III-30 is a similar diagram for the six-channel case. It is representative of the situation for an even number of channels. Note that the integers $\pm M/2$ are combined into a single state. Each time the filter update is performed, the sum $\sum_{i=1}^{M} a_i(j)$ shifts by the integer L_j before the error-correction procedure. After the error-correction procedure, the combined

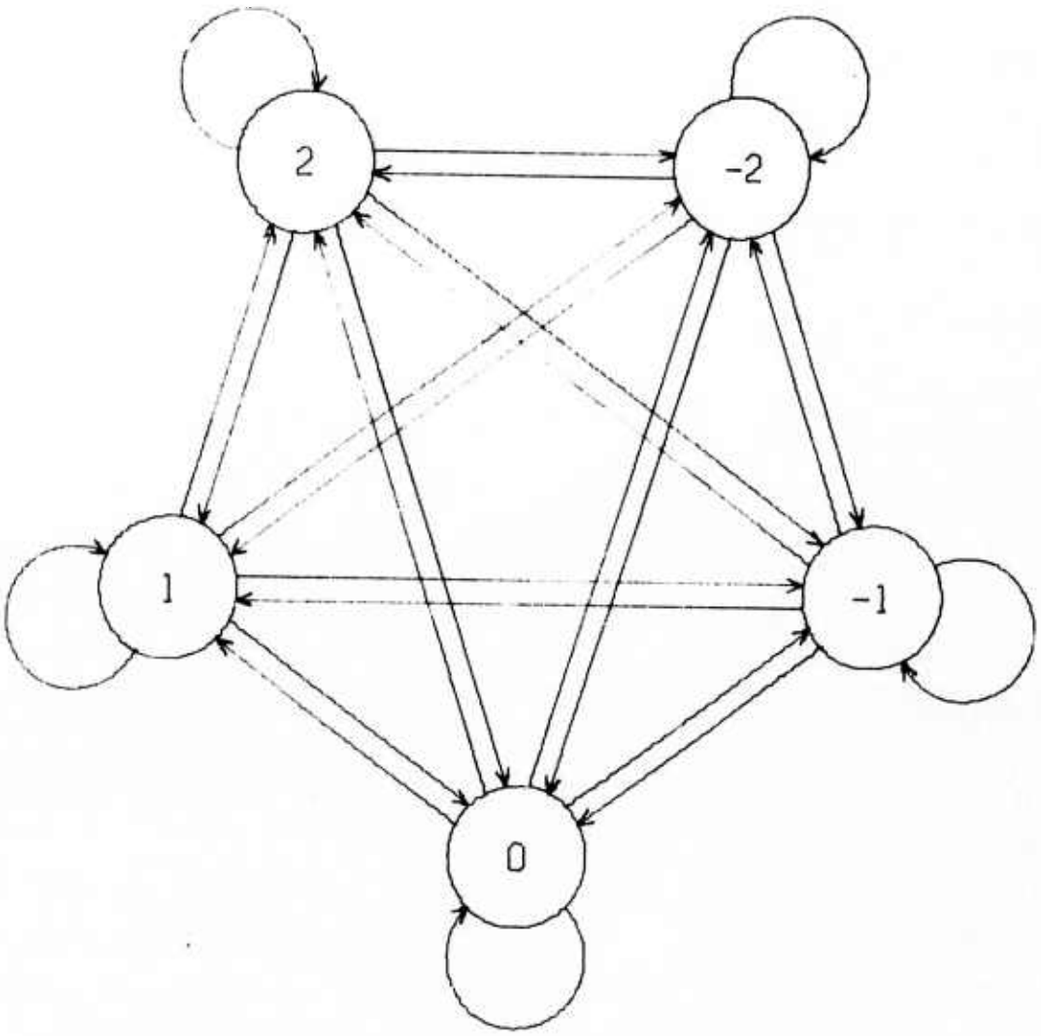


FIGURE III-29

SET OF POSSIBLE INTEGER CONSTRAINT-CONDITION DISCREPANCIES AND TRANSITIONS BETWEEN DISCREPANCIES FOR THE FIVE-CHANNEL CASE

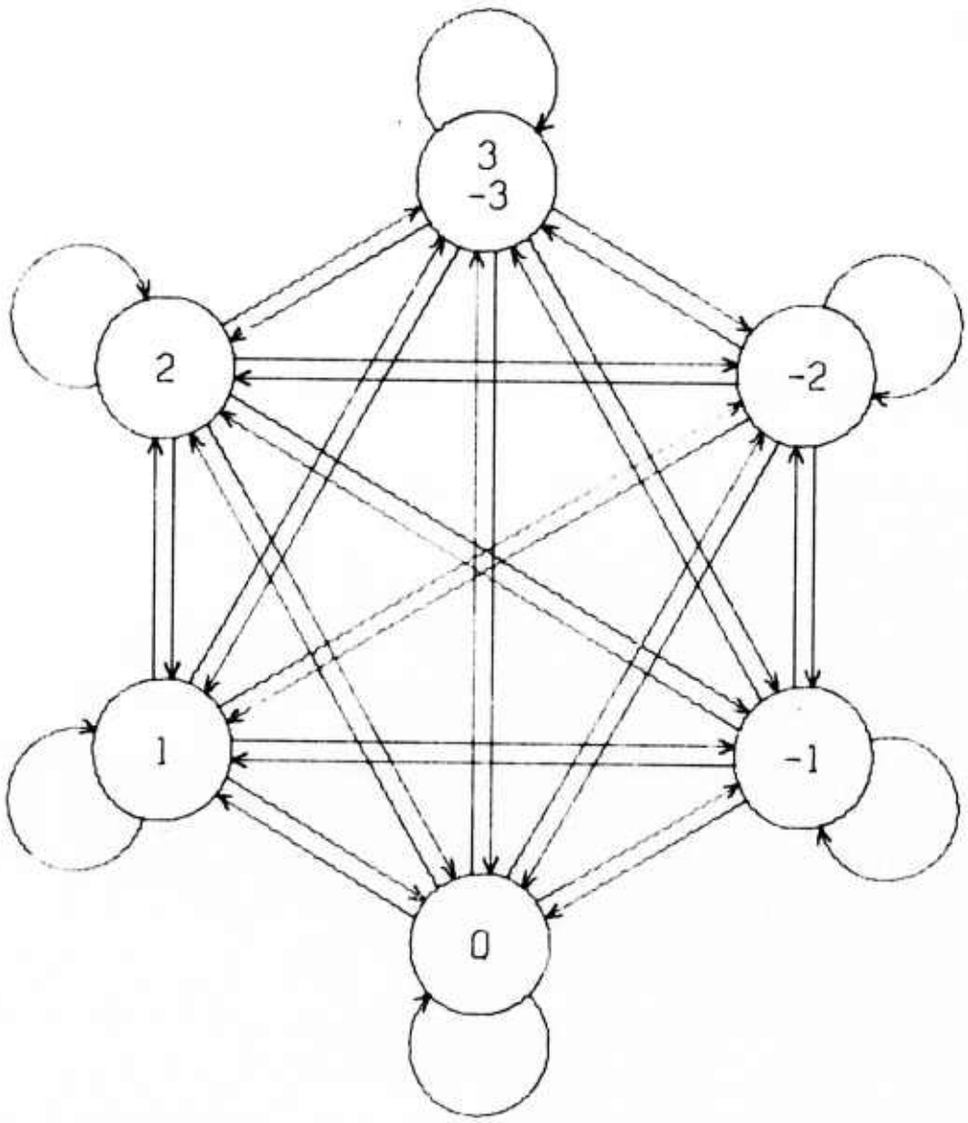


FIGURE III-30

SET OF POSSIBLE INTEGER CONSTRAINT-CONDITION DISCREPANCIES AND TRANSITIONS BETWEEN DISCREPANCIES FOR THE SIX-CHANNEL CASE

effect of the integer shift and correction is a clockwise trip of L_j steps along the outer ring. In the case of an even number of channels, the symmetric probability distribution of L_j about zero guarantees that the integers $-M/2$ and $M/2$ are equally likely after any number of filter updates. The symmetry of L_j about zero also means that the two transition probabilities in the transition matrix corresponding to any pair of states are identical: the Markov transition matrix is symmetric. Since the probability density function for $\sum_{i=1}^{M-1} \delta_i(j)$ is non-zero over the interval from $-(M-1)/2$ to $(M-1)/2$, each of the integers $L_j = -1$, $L_j = 0$, and $L_j = 1$ is always possible whenever $M \geq 3$. As a result, all states are possible at any time after $(M-1)/2$ or $M/2$ updates. When, as in this case, the probability transition matrix for some specified number of updates has no zero elements, the transition probabilities $p_{k\ell}$ into the ℓ-th state for an indefinitely large number of updates approach a fixed limit which is the same for all possible states k (Gnedenko, 1962). Consequently, the probability of any state approaches a limit as the number of updates increases indefinitely, so that the probability of any state is practically independent of any situation in the remote past. Since the probability distribution for the integers L_j does not depend upon the discrepancy between the filter weights and the constraint conditions, any rotation of the state labels along the outer ring does not change the probability transition matrix for any specified number of updates. This fact means that all states are equally likely in the limit as the number of updates approaches infinity. For an odd number of channels, the limiting probability for each possible integer discrepancy is $1/M$. For an even number of channels, it is $1/M$ for each integer discrepancy with absolute value less than or equal to $M/2-1$, while it is $1/2M$ for discrepancies of $-M/2$ and $M/2$. The variance of this probability distribution about its zero mean is $(M^2 - 1)/12$ for odd integers, $(M^2 + 2)/12$ for even integers. When M is 6, the standard deviation from the constraint plane at the j-th lag is

1.77951 filter-weight counts. At the slight cost of allowing this non-cumulative error perpendicular to the constraint space, the mean angle of error $\bar{\theta}$ within the constraint space can be reduced to the level estimated in Figure III-27.

3. Elimination of the Filter-Weight Roundoff Error

The directional error due to the filter-weight roundoff error vector D and the beamsteer roundoff error vector E_b can be eliminated if the adaptive-filter update equation is implemented in the form

$$A^{new} - A^{old} = \left[\frac{2K_s\, y(t)}{M(\bar{X}-X)^T(\bar{X}-X)} \right] M(\bar{X}-X) \quad ,$$

where the factor $2K_s\, y(t)/[M|\bar{X}-X|^2]$ common to all M(2N+1) components is computed as the product of an integer times the quantization level q of the filter weights. When the common factor rounds to zero for non-zero values of y(t), however, the filter vector does not move. To prevent this phenomenon from happening frequently, the quantization level q would need to be chosen so that $2K_s q/\{M \cdot E[|\bar{X}-X|^2]\}$ was close to 1. When $K_s = 0.005$ and $E[|\bar{X}-X|^2] = 9693 \text{ counts}^2$, the required quantization level is $1/5,815,800 \approx 2^{-22}$. Thus for the noise sample from day 238 of 1970 (where the RMS adaptive-filter output was only 3 counts), 22 bits would be needed in the filter-weight representation. To achieve the same goal for data scaled by 16, 30 bits would be needed. It is obviously more practical to perform the complete computation at each component before carrying out any roundoff procedure.

F. SUMMARY

When the data points transmitted from Alaska are multiplied by 16 before processing, the directional error in the vector $(\bar{X}-X)$ and the relative error of the adaptive filter output $y(t) = X^T A$ are reduced to the point where they are virtually the same as the corresponding errors for floating-point arithmetic. Errors in $(\bar{X}-X)$ and $y(t)$ are then limited only by digitization error. The gain doubling performed on the ALPA seismometers in summer of 1972 should permit these errors to be halved.

At the convergence rate $K_s = 0.005$ most frequently used in processing data for this report, the filter-weight roundoff error appears to be the dominant source of error. With the constraint-condition error correction procedure actually employed, the mean angle of error in the filter update vector $(A^{new} - A^{old})$ was approximately 18^o at $K_s = 0.005$ (ignoring sign reversals in the adaptive filter output). If the alternate constraint-condition error correction procedure (described in part 2 of Subsection E) had been implemented, the mean angle of error in $(A^{new} - A^{old})$ would have been approximately 10^o. At convergence rates below $K_s = 0.005$, the filter-weight roundoff error seems large enough to affect adaptive-filter noise reduction relative to beamsteering (see Figures IV-1 and IV-2 in Section IV).

SECTION IV

SIGNAL-TO-NOISE IMPROVEMENT IN THE PRESENCE OF BACKGROUND NOISE

A. INTRODUCTION

In this section, the subject of investigation is the ability of adaptive multichannel filtering to provide signal-to-noise ratio improvement relative to beamsteering in the presence of background noise. The critical area of concern is the improvement for weak signals in the borderline detection range (where the signal-to-noise ratio in the beamsteer output is between 6 and 12 dB). In this range of signal-to-noise ratios, detection procedures are unable to recognize the presence of a signal with any consistency. Under these circumstances it is not possible to rely on the standard procedure of freezing an adaptive multichannel filter set when a signal is detected. Results in this section are obtained solely from computer runs that implement the adaptive filtering process at each point in time when a new data sample is available. One signal approximately 6 dB above the noise level on the beamsteer output is used for the critical results. Two other signals with signal-to-noise ratios of 18 and 24 dB are processed in this way to illustrate the desirability of preventing the adaptive filter update in the presence of a signal. Provided that a signal of this size is not overlapped by an interfering event, any dire effects on the signal due to adaptive filtering are of no serious concern: the signal-to-noise ratio on the beamsteer output is sufficiently high that the beamsteer output is adequate for subsequent analysis of the event.

The method of approach used to estimate signal-to-noise ratio improvement is to measure the difference between noise reduction and signal degradation. Both noise reduction and signal degradation are computed in terms of the

drop in power for the adaptive-filter output relative to the beamsteer output. Since it is impossible in the case of real data to distinguish between signal power and noise power when a signal is present, the data samples for signal measurements must be different from those used for noise measurements. Noise reduction is computed over data samples approximately four hours long. Signal degradation, on the other hand, is computed over four- to eight-minute gates. In the case of signal degradation measurements, the effect of adaptive filtering on contaminating noise is lumped together with the effect on the signal proper. As long as the adaptive filter set reduces the contaminating noise power at least as much as the signal power, the signal degradation is not underestimated and the signal-to-noise ratio improvement is not overestimated. Noise reduction and signal degradation are calculated for the same pair of data samples over a range of convergence rates. From these figures, the signal-to-noise ratio is determined as a function of convergence rate.

At convergence rates where signal distortion and reliability of the noise statistics are not serious problems, there are two principal factors determining adaptive-filter performance. As the convergence rate increases, signal degradation climbs to a damaging level. Conversely, as the convergence rate decreases, the ability of the adaptive filter to respond to changes in the noise field is impaired. The convergence rate at which signal-to-noise ratio improvement is maximized lies somewhere in between.

Subsection B discusses noise reduction as a function of convergence rate. Subsection C presents noise reduction as a function of frequency for one convergence rate close to optimum. Subsection D studies signal degradation as a function of convergence rate. Finally, the noise reduction and signal degradation results are combined in Subsection E to yield signal-to-noise ratio improvement as a function of convergence rate.

B. NOISE REDUCTION AS A FUNCTION OF CONVERGENCE RATE

Two samples, one from day 238 of 1970 and the other from day 203 of 1971, were used to investigate the effect of convergence rate on adaptive filter noise reduction relative to beamsteering. The look direction for both samples corresponds to an azimuth of 270°. Noise reduction is calculated as the ratio of beamsteer output power to adaptive-filter output power. It is computed for the entire band 0.0 to 0.5 Hz and also for the band 0.0234 to 0.0664 Hz (corresponding to periods between 43 and 15 seconds). Each point on the noise reduction curves shown in Figures IV-1 and IV-2 was generated by passing approximately four hours of data through the adaptive filtering package at a particular convergence rate. The noise reduction shown is for the entire length of the data samples processed. Convergence rates are defined as percentages of the convergence rate $K_s = 1$.

The first data sample to be run through the adaptive filtering package covers the period 0757 to 1150 on day 238 of 1970. The vertical components of sites 1, 2, 3, 6, 8, and 9 from the ALPA array were used as input channels for the adaptive-filter beam. The PDE bulletin lists no events between 0639 and 1502 on August 26 of 1970. A scan of the four-hour sample from 12° to 348° at 24° increments using the Fisher detector (see Subsection V-A) indicates, however, that a signal reached ALPA at 1021 from an azimuth just under 60°, possibly from the North Atlantic ridge. (Beam traces steered for 60° show a frequency-dispersed Love-wave arrival prior to frequency-dispersed Rayleigh-wave arrivals on the vertical and radial components.) The energy from this signal is only a small fraction of the total energy within the four-hour sample and should have only a minor effect on the reported noise reduction values. Table IV-1 shows noise reduction in dB as a function of convergence rate both broadband and in the band corresponding to periods between 43 and 15 seconds. Figure IV-1 is a graph of noise reduction given in Table IV-1. Maximum broadband noise reduction was achieved near a convergence rate of 35%. This

TABLE IV-1

ADAPTIVE FILTERING NOISE REDUCTION VERSUS CONVERGENCE RATE (DAY 238 1970)

Convergence Rate (% of maximum)	Broadband Noise Reduction (dB)	Narrowband Noise Reduction (dB)
0.2	1.099	1.156
0.3	1.359	1.439
0.4	1.521	1.618
0.5	1.639	1.746
0.7	1.813	1.936
1.0	1.983	2.124
1.5	2.186	2.347
2.0	2.346	2.526
3.0	2.635	2.847
5.0	3.135	3.415
7.0	3.562	3.911
10.0	4.056	4.501
15.0	4.547	5.108
20.0	4.786	5.413
25.0	4.892	5.563
30.0	4.940	5.639
35.0	4.945	5.666
40.0	4.939	5.667
50.0	4.896	5.670

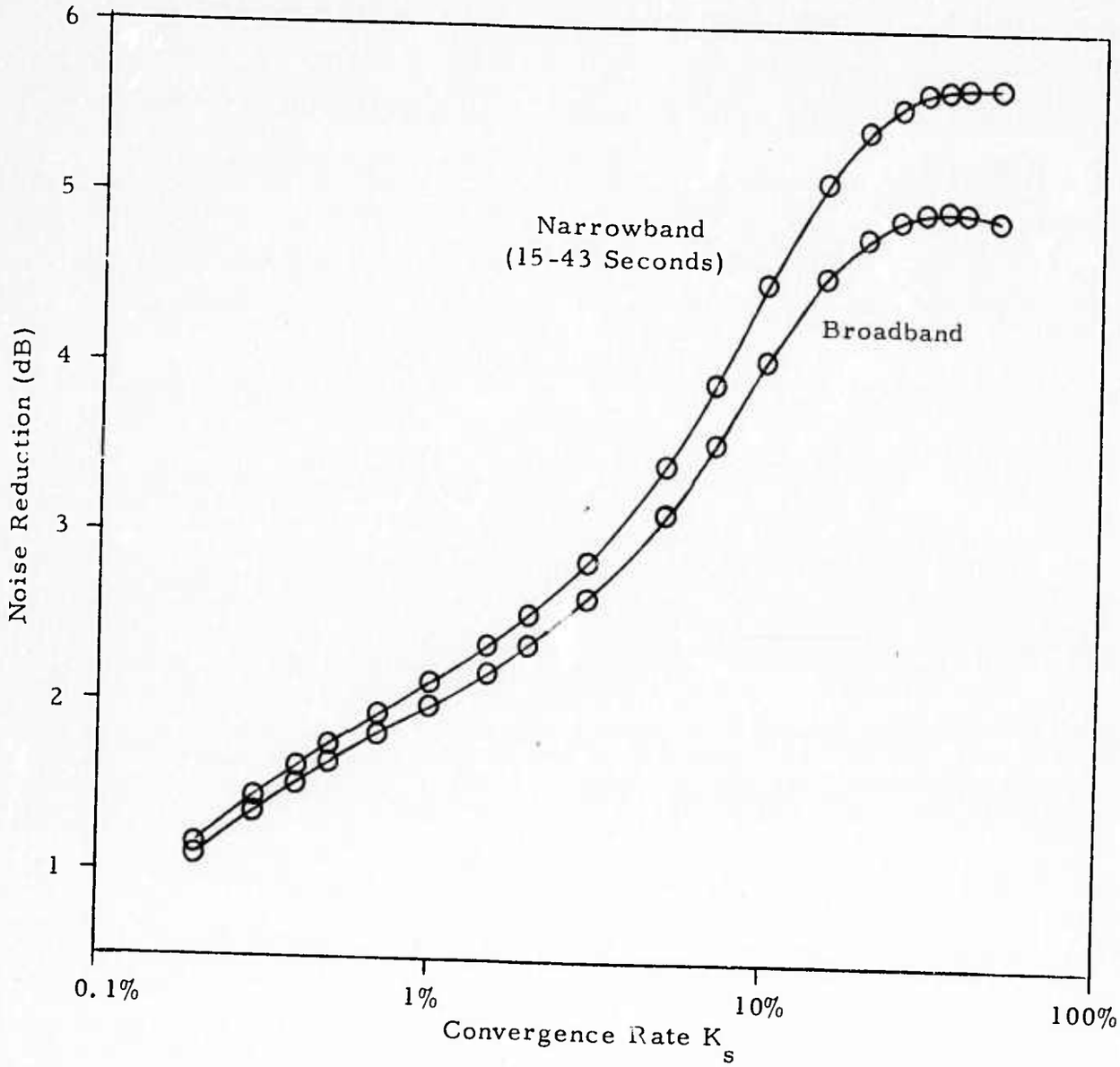


FIGURE IV-1

NOISE REDUCTION AS A FUNCTION OF CONVERGENCE RATE
DAY 238 OF 1970 (0757 TO 1150 GMT)

fact is probably of little importance, since the correlation between one data vector X (see page III-1 for a definition) and the next is high. The high correlation occurs because the highest frequency (0.0625 - 0.125 Hz) at which there is any significant noise is about one quarter to one eighth of the Nyquist frequency associated with a one-second sample interval. If one data vector X were exactly the same as the next, for example, the noise power would be reduced to $(1 - 2K_s)^2$ of its original value: the update equation

$$A^{new} = A^{old} + \frac{2K_s(\bar{X}-X)\,X^T A^{old}}{(\bar{X}-X)^T(\bar{X}-X)}$$

would imply that

$$X^T A^{new} = X^T A^{old}\left[1 + \frac{2K_s X^T(\bar{X}-X)}{(\bar{X}-X)^T(\bar{X}-X)}\right]$$

$$= X^T A^{old}\left[1 - \frac{2K_s(\bar{X}-X)^T(\bar{X}-X)}{(\bar{X}-X)^T(\bar{X}-X)}\right]$$

$$= X^T A^{old}(1 - 2K_s)$$

and

$$(A^T)^{new} X X^T A^{new} = (A^T)^{old} X X^T A^{old}(1 - 2K_s)^2 \quad .$$

This result would also apply to data vectors X containing signals. And so it is necessary to examine what happens to signals at corresponding convergence rates.

The second data sample encompasses the interval 0355 to 0725 on day 203 of 1971. Sites 8, 13, 14, 15, 16, and 17 of the ALPA array were input to the adaptive-filter beam. The PDE bulletin reports some activity in the New Ireland and Northern Celebes regions of the South Pacific during this period (notably a magnitude 5.4 event from 0.4N, 123.5E at 06:07:52.8 GMT). Since the azimuth from which this energy arrives is extremely close to 270° (the look direction), the data sample contains signal energy. The relative strength of energy from seismic events in this data sample is discussed later in Subsection B of Section VI. Apparently a sizeable fraction of the total energy is due to seismic events (both on-azimuth and off-azimuth). Table IV-2 and Figure IV-2 give the noise reduction in dB for this sample as a function of convergence rate in the frequency bands 0.0-0.5 Hz and 0.0234-0.0664 Hz. Maximum broadband noise reduction occurs around a 25-30% convergence rate. Note that the broadband noise reduction drops abruptly between a 32% and 35% convergence rate, whereas the narrowband reduction stays more or less level all the way to a 50% convergence rate. The reason for this drop is that the adaptive filter set temporarily diverged over a short section of data containing a glitch. The result of the divergence was a rapid oscillation in the adaptive filter output at the folding frequency. The folding frequency does not show up in the narrowband analysis, and hence the narrowband noise reduction remains relatively constant beyond a 32% convergence rate.

Note that the noise reduction in both these samples begins to arc downward slightly as the convergence rate decreases from $K_s = 0.5\%$ to $K_s = 0.2\%$. The effects of rounding the filter weights to the nearest computer count are probably becoming evident. A less plausible explanation is that the adaptive filter set utilizes a sizeable fraction of the total four-hour sample before it completes the transition from the initial beamsteer weights to a filter set reflecting the statistics of the noise field.

TABLE IV-2

ADAPTIVE FILTERING NOISE REDUCTION VERSUS CONVERGENCE RATE (DAY 203 1971)

Convergence Rate (% of maximum)	Broadband Noise Reduction (dB)	Narrowband Noise Reduction (dB)
0.2	1.705	1.828
0.3	2.044	2.181
0.4	2.275	2.421
0.5	2.442	2.596
0.7	2.686	2.852
1.0	2.937	3.123
1.5	3.207	3.429
2.0	3.420	3.679
3.0	3.782	4.122
5.0	4.362	4.826
7.0	4.793	5.335
10.0	5.246	5.864
15.0	5.611	6.293
20.0	5.764	6.481
25.0	5.814	6.547
30.0	5.807	6.561
32.0	5.806	6.562
33.0	5.732	6.557
35.0	5.494	6.558
40.0	5.393	6.537
50.0	5.281	6.498

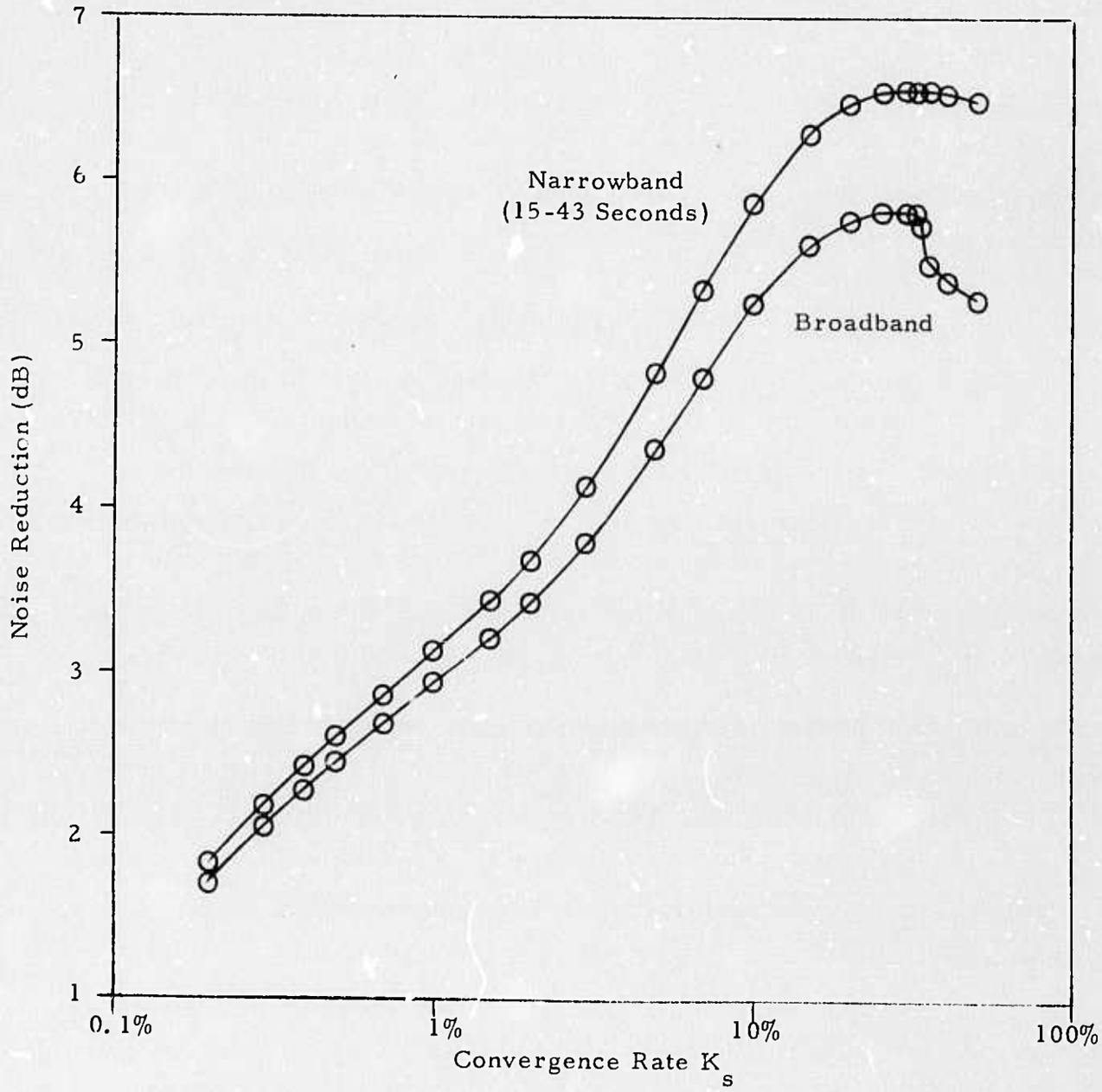


FIGURE IV-2

NOISE REDUCTION AS A FUNCTION OF CONVERGENCE RATE
DAY 203 OF 1971 (0355 TO 0725 GMT)

Both of the noise samples studies so far were taken from summertime (when the noise is at its lowest). In Subsection C, noise reduction as a function of frequency for a sample from day 7 of 1972 is discussed briefly. Also in Section VI, where noise reduction is examined as a function of steer direction, a data sample near the peak autumn noise level is processed. These two samples, one from winter and one from fall, give some idea of the reduction of background noise achievable later in the year.

C. NOISE REDUCTION AS A FUNCTION OF FREQUENCY

To determine the frequencies at which adaptive filtering was producing noise reduction, power density spectra for the two beams were computed. Spectra in this report are given in dB relative to one millimicron squared per Hz at frequencies where the prefilter response is unity. If square-root-of-N noise reduction had been achieved through beamsteering, equivalent single-sensor noise levels would be 7.8 dB (i.e., $10\log_{10}6$) higher than the beamsteer power density spectra.

Figure IV-3 presents output spectra of the conventional and adaptive-filter beams for the noise sample from day 238 of 1970. The spectrum rises to a peak of 26.5 dB at a period of 18 seconds and descends rapidly from this peak as the frequency increases. Figure IV-4 is a graph of the noise reduction from adaptive filtering at a convergence rate 0.5% of maximum. Noise reduction is measured as the ratio of the beamsteer output power density to the adaptive-filter output power density. Peak noise reduction of 3.0 dB is obtained at a frequency corresponding to a period of 20 seconds. Noise reduction of 1 dB is maintained at almost all periods above 13 seconds. Positive noise reduction is preserved at all frequencies where there is significant power.

Corresponding beam output spectra for the noise sample from day 203 of 1971 are plotted in Figure IV-5. The adaptive-filter convergence rate was

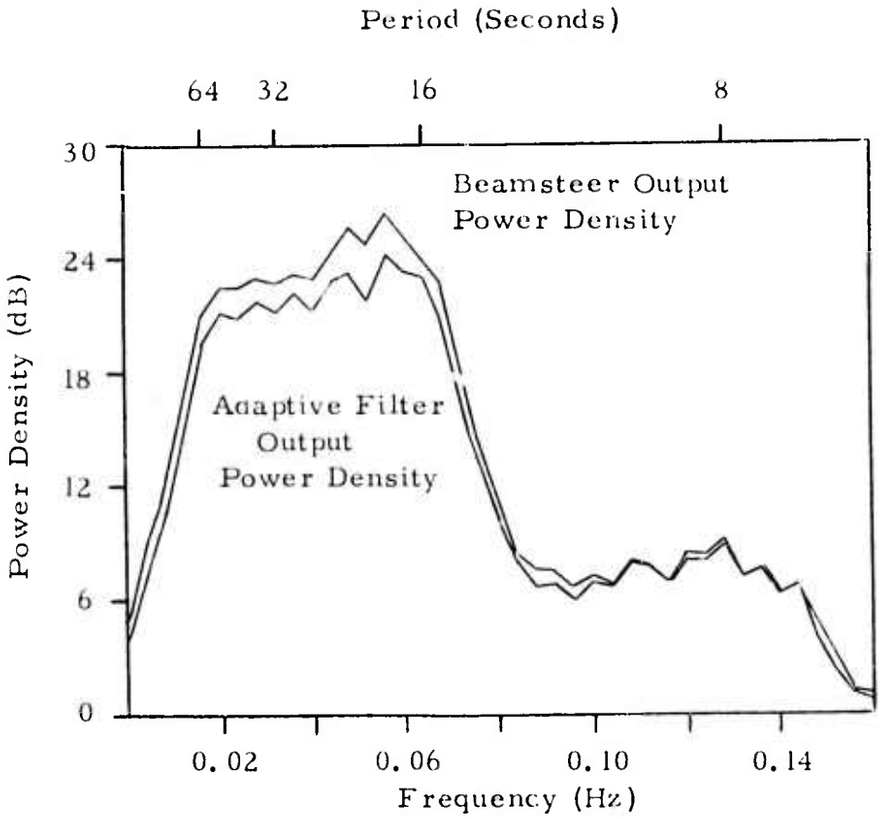


FIGURE IV-3

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
DAY 238 1970, STEER DIRECTION 270°, K_s = 0.005)

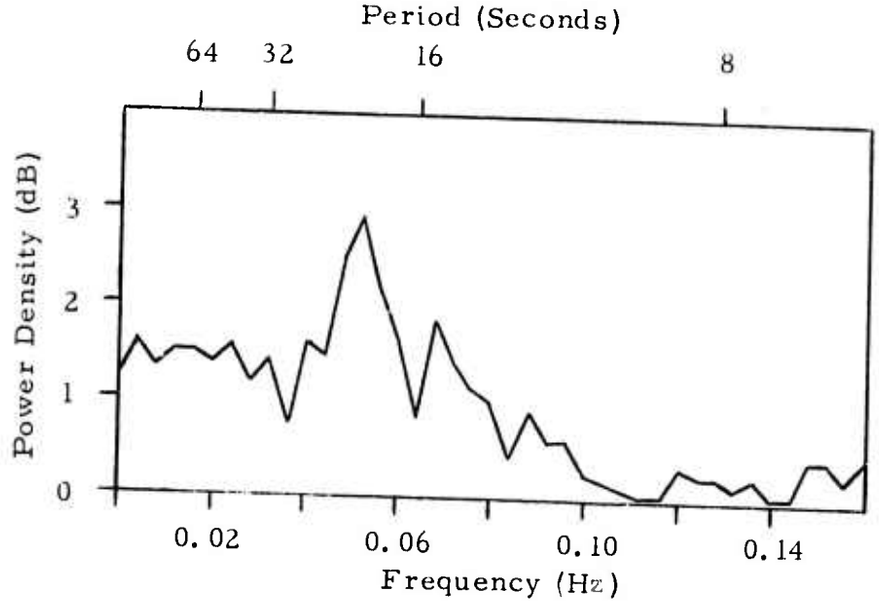


FIGURE IV-4

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY (DAY 238 1970, STEER DIRECTION 270°, $K_s = 0.005$)

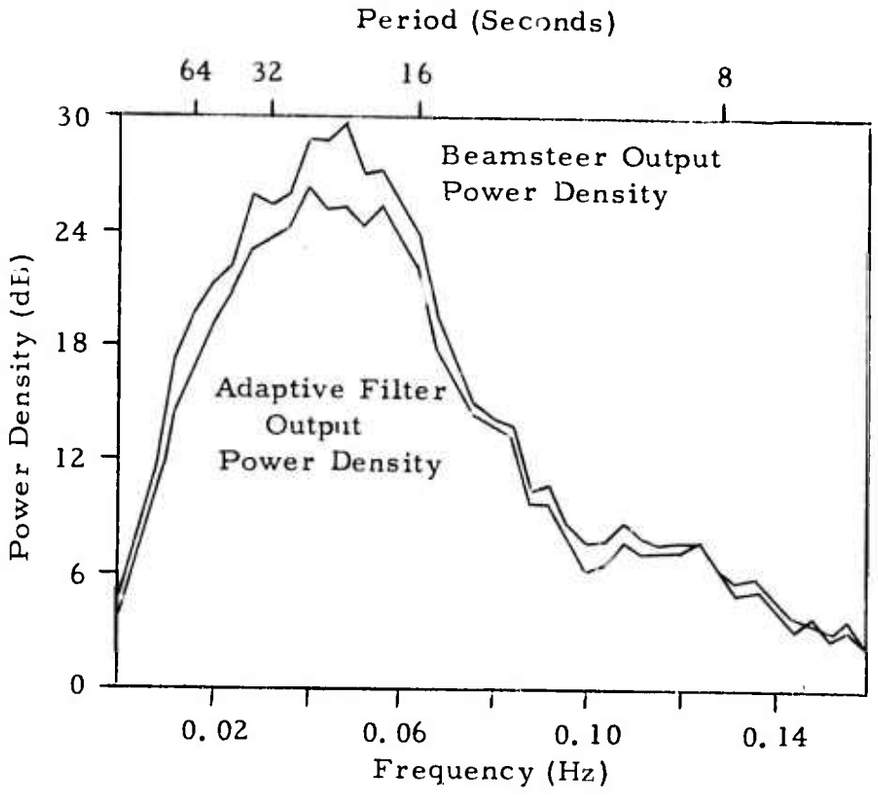


FIGURE IV-5

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 203 1971, STEER DIRECTION 270°, K_s = 0.005)

again 0.5% of maximum. The beamsteer spectrum rises to a peak with periods in the range 21-26 seconds. At higher frequencies, spectral level drops significantly. Figure IV-6 gives the adaptive-filtering noise reduction relative to beamsteering. In this sample, peak noise reduction of 4.3 dB occurs at a period of 21 seconds. Once again, positive noise reduction is maintained at all frequencies with significant power. Due to the presence of a glitch and on-azimuth events in the data, some caution should be exercised in the interpretation of the results (particularly in connection with the three major peaks in noise reduction). Possibly the reduction in the noise proper is comparable to that for day 238 of 1970.

Another noise sample from day 7 of 1972 was processed. It was atypical because the spectrum peaked at 7 1/2 seconds: the principal microseismic peak at 17 seconds was 28.5 dB relative to one millimicron squared per Hz, whereas the secondary microseismic peak was 32 dB at 7 1/2 seconds. Highest noise reduction was 3.5 dB at 17 1/2 seconds. In contrast, the noise reduction at 7 1/2 seconds was only 2.2 dB. At a convergence rate 0.5% of maximum, broadband noise reduction was 1.8 dB, narrowband noise reduction 2.1 dB.

Evidently noise reduction is greatest at the principal spectral peaks of the ALPA noise. After the prefilter of Figure II-1 is applied to the vertical-component data, these peaks are (1) the principal microseismic peak near 18 seconds, and (2) the secondary microseismic peak near 8 seconds.

D. SIGNAL DEGRADATION AS A FUNCTION OF CONVERGENCE RATE

To probe the effects of convergence rate on adaptive-filter signal degradation, three signals were selected out of a four-hour sample from day 276 of 1971 (October 3). The sample spans the time period 2000 to 2357. Sites 8, 9, 12, 15, 16, and 17 are input to the time-shift-and-sum and adaptive-filter beams. In this analysis, the adaptive-filter is permitted to update under all conditions and is never frozen.

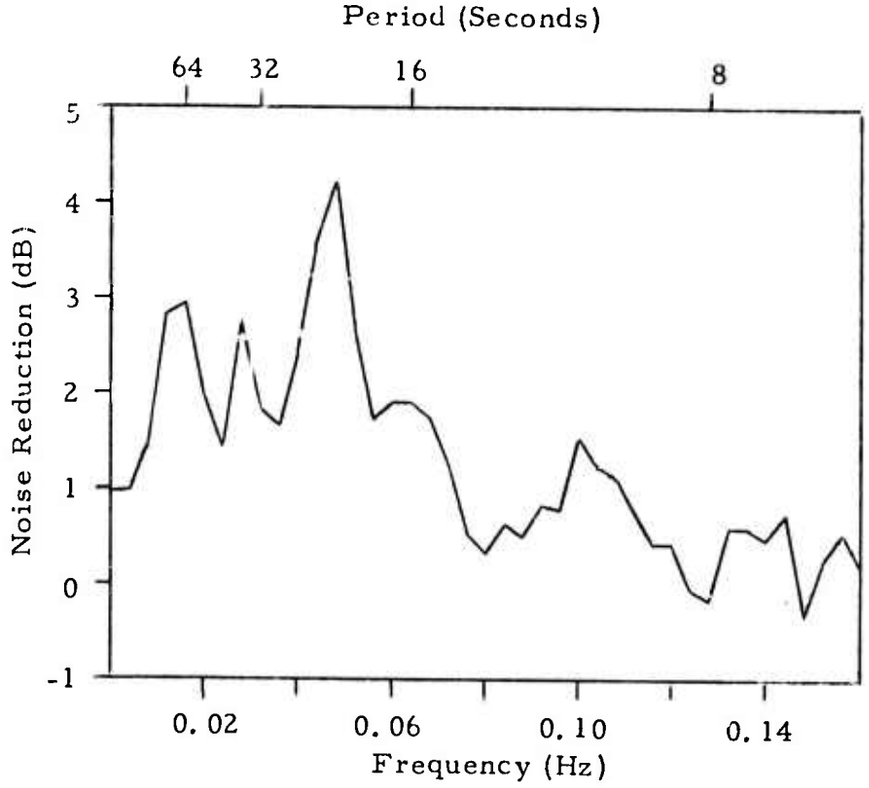


FIGURE IV-6

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY (DAY 203 1971, STEER DIRECTION 270°, K_s = 0.005)

The critical signal arrived at ALPA at approximatel/ 2140. The signal-to-noise ratio on the beamsteer output is about 6 dB. No event corresponding to this signal can be found on either the PDE bulletin or LASA bulletin. Various forms of frequency-wavenumber spectra computed at several frequencies peak at azimuths ranging from 299° to 306°. A Fisher-detector scan of this signal, moreover, shows a higher detector output at 300° than at the two neighboring look azimuths of 276° and 324°. The beamsteer and adaptive-filter beams are aimed toward 302.5°. Signal degradation for this event is computed at convergence rates from 0.2% to 25% of maximum over a 512-second gate. Table IV-3 and Figure IV-7 depict signal degradation as a function of convergence rate both broadband and in the frequency band associated with periods between 43 and 15 seconds. As the convergence rate increases beyond 0.5%, the narrowband degradation rises above the broadband degradation. The negative degradation at the 0.2% rate means that the adaptive filter trace contains more energy than the beamsteer output trace over the 512-point gate. (As a check on the measurement technique, the convergence rate was set equal to zero, and the result was a power difference no larger than 0.005 dB.) Figure IV-8 shows the beamsteer and adaptive-filter output for this event at a convergence rate of 0.5%. Tick marks at the zero levels of both beams are spaced 5 minutes apart and correspond to the times given midway between the traces. The beamsteer output is shown in the top trace, the adaptive-filter output in the bottom trace.

The first of the two large signals examined in this section is an event listed on the PDE bulletin. This earthquake from near the east coast of Kamchatka has a bodywave magnitude of 4.5. The origin time is 20:54:48.5 (again on day 276 of 1971), the latitude 55.7 N, the longitude 162.1 E. The azimuth with respect to ALPA is 273°. Accordingly, the beams are trained in this direction. 18 dB is the signal-to-noise ratio on the beamsteer output. Signal degradation is computed over a 256-second gate at convergence rates

TABLE IV-3

ADAPTIVE FILTERING SIGNAL DEGRADATION VERSUS CONVERGENCE RATE FOR A WEAK SIGNAL FROM 300°-305° (DAY 276 1971)

Convergence Rate (% of maximum)	Broadband Signal Degradation (dB)	Narrowband Signal Degradation (dB)
0.2	-0.024	-0.039
0.3	0.104	0.093
0.4	0.255	0.251
0.5	0.408	0.408
0.7	0.650	0.655
1.0	0.945	0.956
1.5	1.301	1.320
2.0	1.569	1.599
3.0	1.970	2.031
5.0	2.562	2.686
7.0	3.049	3.229
10.0	3.713	3.948
15.0	4.505	4.781
20.0	4.973	5.283
25.0	5.176	5.526

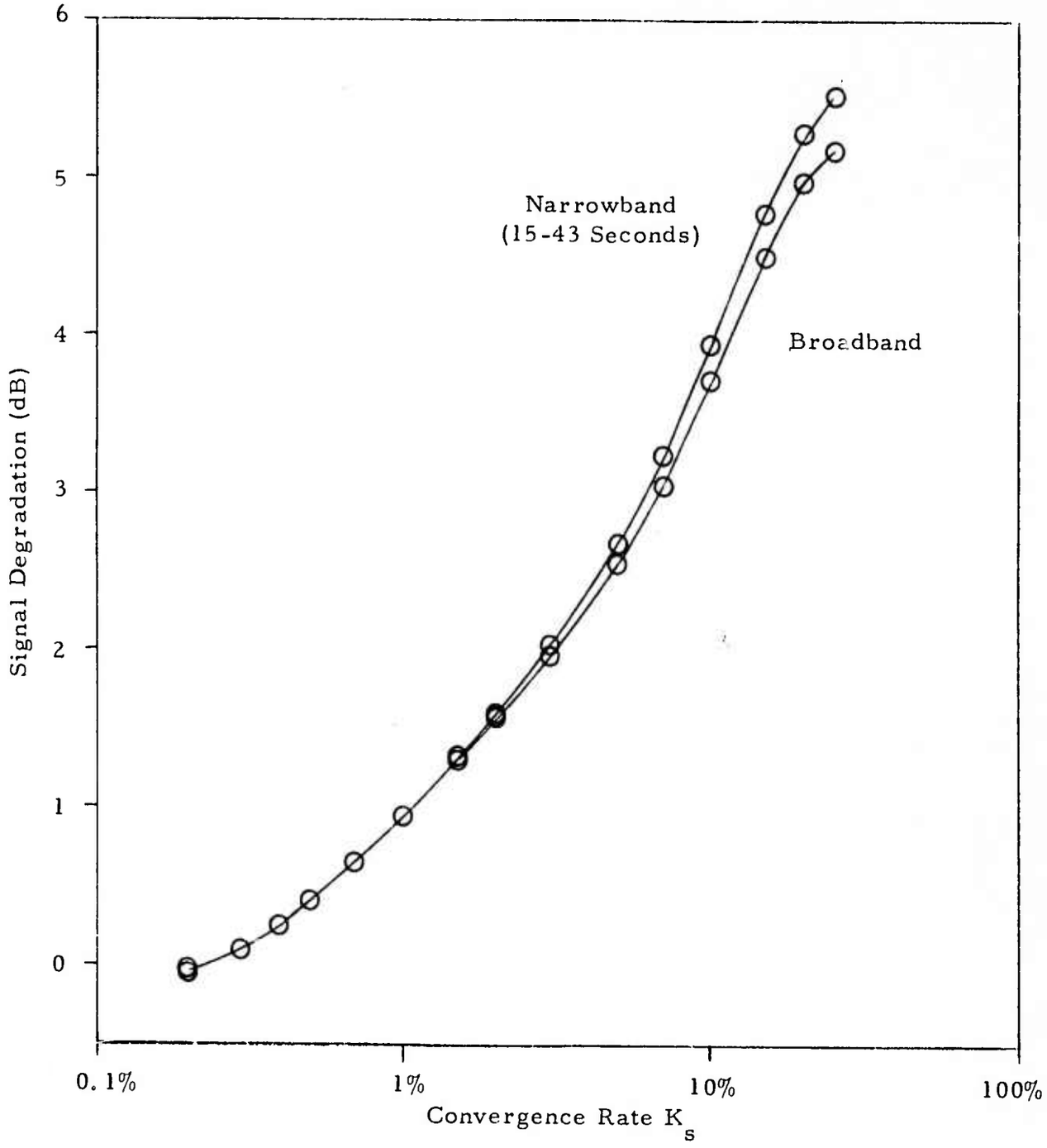


FIGURE IV-7

SIGNAL DEGRADATION AS A FUNCTION OF CONVERGENCE RATE FOR A WEAK SIGNAL APPROXIMATELY 6 dB ABOVE NOISE LEVEL ON BEAMSTEER OUTPUT (STEER AZIMUTH 302.5°)

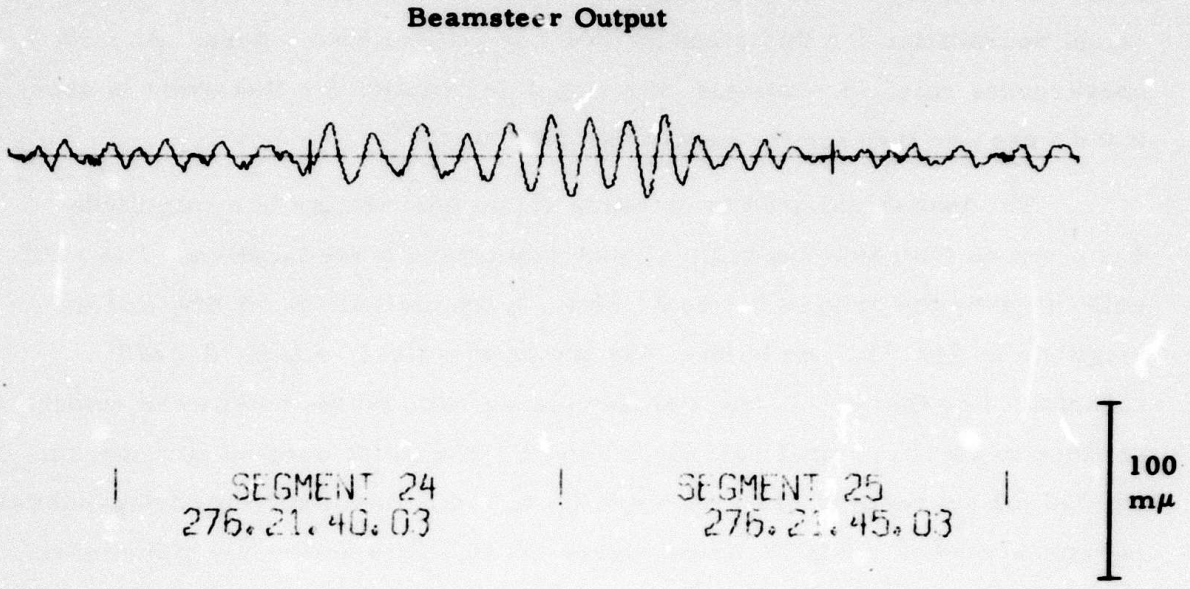


FIGURE IV-8

WEAK EVENT FROM 300°-305°
(FILTER ADAPTING, STEER DIRECTION 302.5°, K_s = 0.005)

varying from 0.2% to 25%. The resulting signal degradation figures are portrayed in Table IV-4 and Figure IV-9. For convergence rates up to 10%, the signal degradation for this stronger event is greater than before. At a 5% convergence rate, for example, the signal degradation for this event is about 0.9 dB greater than for the event from 300^{o} to 305^{o}.

The last signal used to measure signal degradation is a magnitude 4.9 event on the same day from almost exactly the same location. The PDE bulletin gives the origin time as 21:54:12.9, the latitude as 55.8N, and the longitude as 162.2E. As before, the beams are directed toward a 273^{o} azimuth. For this event, the signal-to-noise ratio on the beamsteer output is close to 24 dB. Signal degradation over a 256-point interval is again calculated for convergence rates from 0.2% to 25%. Resultant signal degradation is represented in Table IV-5 and Figure IV-10. It is noticeably higher than for the event with an 18-dB signal-to-noise ratio. Figure IV-11 pictures the beamsteer output and adaptive-filter output for this tremor when the convergence rate is 0.5%. At that rate, degradation is just under 2 dB. The worst distortion occurs at the end of the event.

Two features in the signal degradation curves for the larger events are worth mentioning. As the signal-to-noise ratio of these events increases, so does the signal degradation. With the particular adaptive algorithm employed, it is abundantly clear that the maximum-likelihood constraints do not guarantee that the signal is preserved. Furthermore, the signal degradation begins to approach a maximum at lower convergence rates than does the noise reduction.

E. SIGNAL-TO-NOISE GAIN AS A FUNCTION OF CONVERGENCE RATE

In this subsection, signal degradation is subtracted from noise reduction to yield signal-to-noise ratio improvement. This procedure is carried out over a range of convergence rates in order to discover the convergence rates where the largest signal-to-noise gain is achieved.

TABLE IV-4

ADAPTIVE FILTERING SIGNAL DEGRADATION VERSUS CONVERGENCE RATE FOR A STRONG SIGNAL FROM KAMCHATKA (DAY 276 1971)

Convergence Rate (% of maximum)	Broadband Signal Degradation (dB)	Narrowband Signal Degradation (dB)
0.2	0.269	0.268
0.3	0.416	0.410
0.4	0.556	0.544
0.5	0.693	0.674
0.7	0.944	0.915
1.0	1.279	1.236
1.5	1.743	1.684
2.0	2.137	2.070
3.0	2.718	2.650
5.0	3.423	3.369
7.0	3.775	3.735
10.0	3.989	3.958
15.0	4.119	4.085
20.0	4.203	4.170
25.0	4.243	4.214

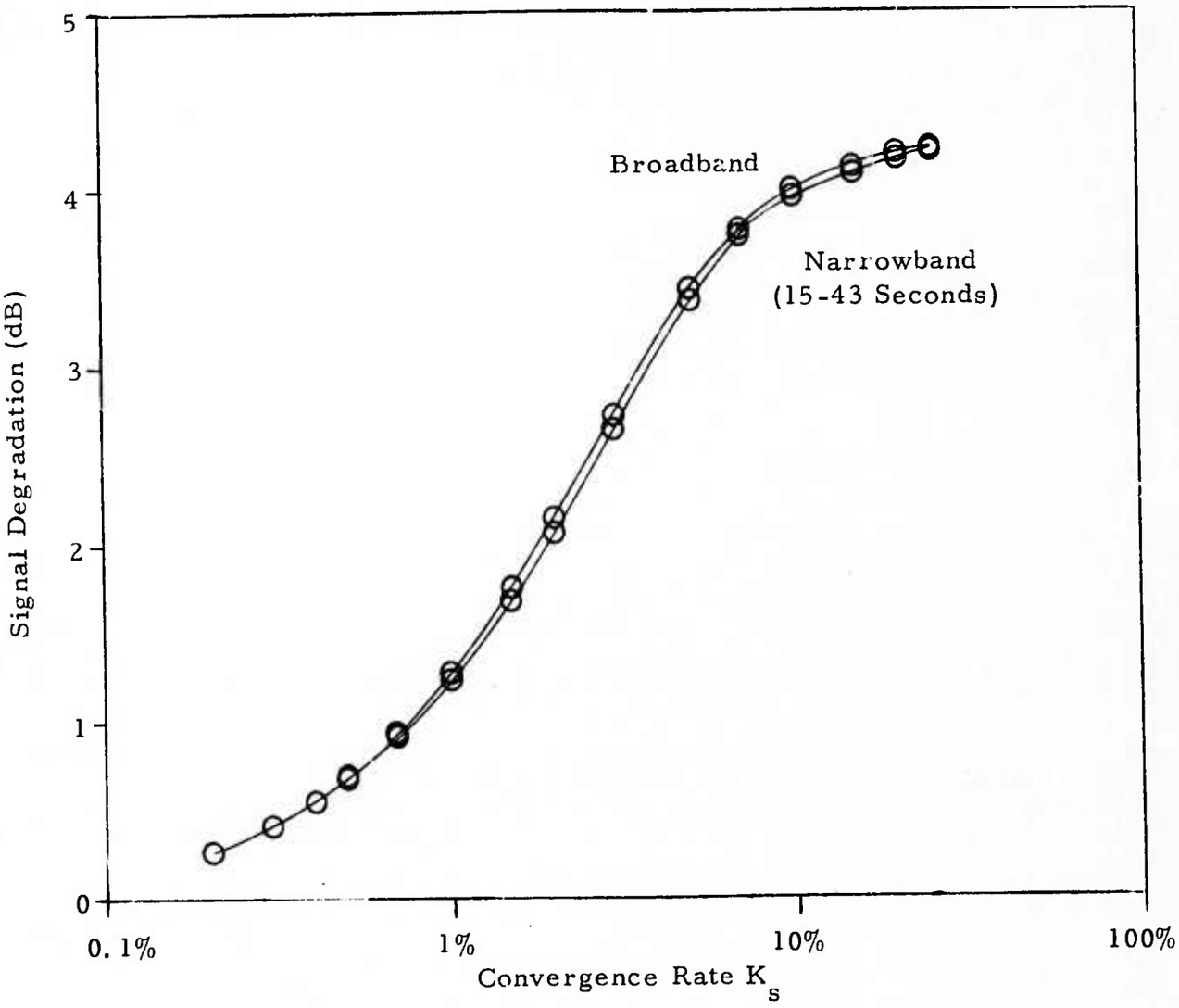


FIGURE IV-9

SIGNAL DEGRADATION AS A FUNCTION OF CONVERGENCE RATE FOR A SIGNAL 18 dB ABOVE NOISE LEVEL ON BEAMSTEER OUTPUT

TABLE IV-5

ADAPTIVE FILTERING SIGNAL DEGRADATION VERSUS CONVERGENCE RATE RATE FOR A VERY STRONG SIGNAL FROM KAMCHATKA (DAY 276 1971)

Convergence Rate (% of maximum)	Broadband Signal Degradation (dB)	Narrowband Signal Degradation (dB)
0.2	0.963	0.948
0.3	1.324	1.305
0.4	1.655	1.633
0.5	1.977	1.954
0.7	2.571	2.550
1.0	3.360	3.353
1.5	4.450	4.482
2.0	5.294	5.371
3.0	6.367	6.524
5.0	7.179	7.417
7.0	7.418	7.693
10.0	7.461	7.764
15.0	7.465	7.800
20.0	7.495	7.850
25.0	7.497	7.862

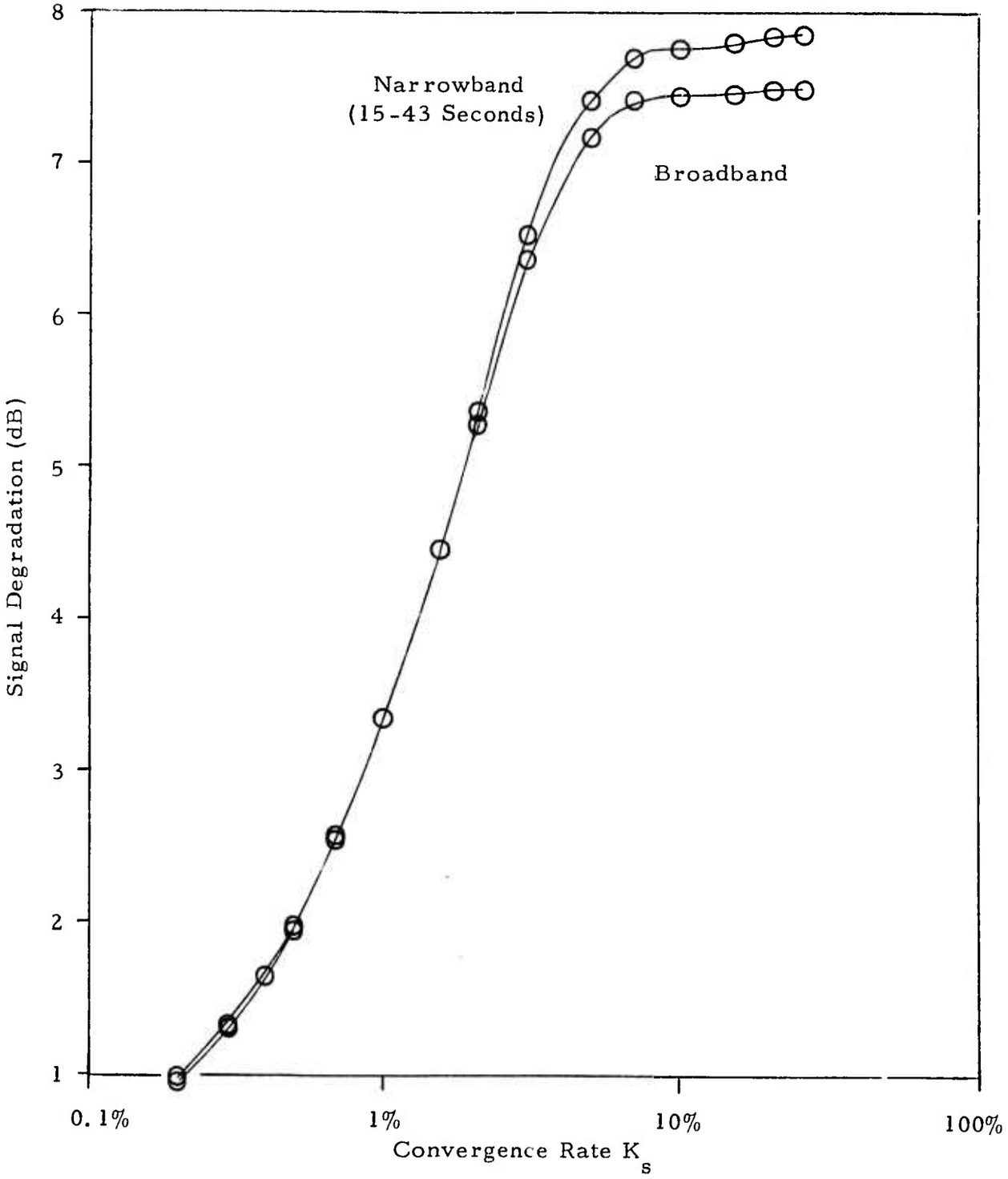


FIGURE IV-10

SIGNAL DEGRADATION AS A FUNCTION OF CONVERGENCE RATE FOR A SIGNAL 24 dB ABOVE NOISE LEVEL ON BEAMSTEER OUTPUT

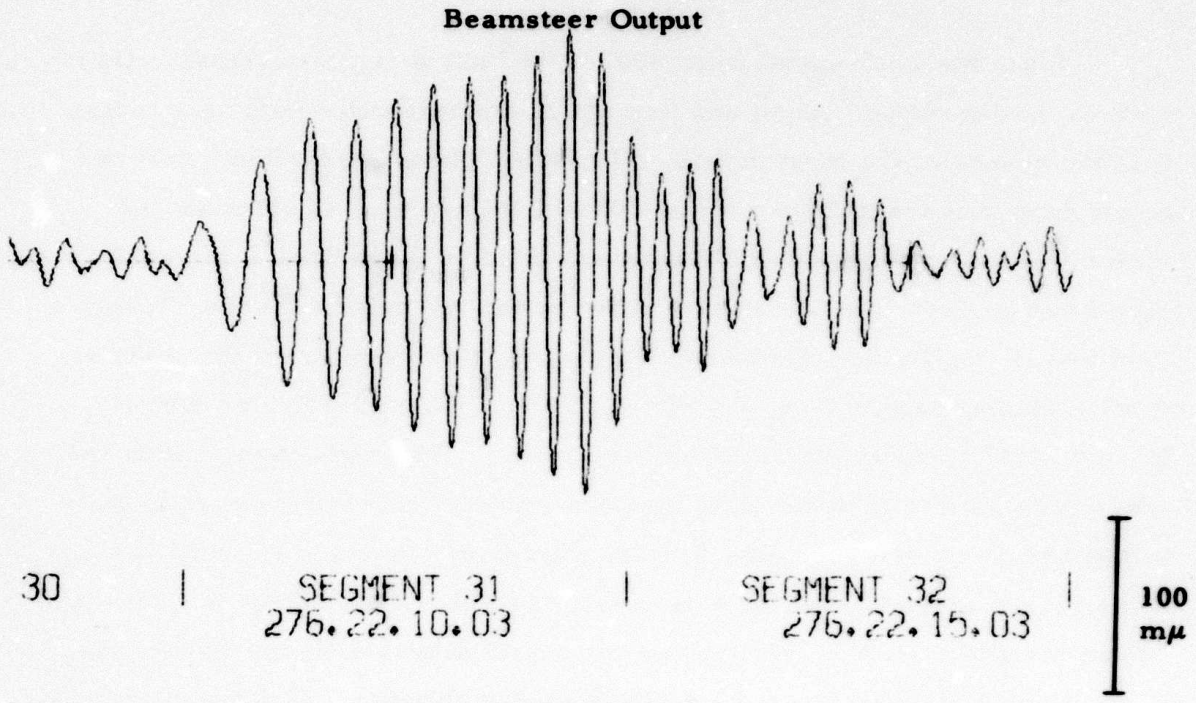


FIGURE IV-11

MAGNITUDE 4.9 EVENT FROM KAMCHATKA (FILTER ADAPTING, STEER DIRECTION 273°, K_s = 0.005)

Since the weak signal from 300° - 305° has a signal-to-noise ratio of 6 dB on the beamsteer output and lies within the borderline detection range, it is the source of the most meaningful results. Noise reduction measurements from day 238 of 1970 and day 203 of 1971 are used to estimate the signal-to-noise gain which would have occurred on those days if the weak signal had arrived at ALPA during either of the two noise samples. Tables IV-6 and IV-7 give the signal-to-noise gain for this event using the noise reduction measurements from day 238 of 1970 and 203 of 1971, respectively. Figure IV-12 contains the results for both days. Narrowband gain is at the top in each pair of curves. Best signal-to-noise ratio improvement is obtained near a convergence rate 0.5% of maximum. At that rate, signal-to-noise gain is about 1.3 dB for noise from day 238 of 1970, about 2.1 dB for "noise" from day 203 of 1971. Since the signal degradation approaches the same order of magnitude as the noise reduction at higher convergence rates, it becomes increasingly difficult to determine what proportion of the remaining adaptive filter output is from the weak signal. The high ratio of noise to signal in this event further complicates the situation at higher convergence rates. As a result, the curves in Figure IV-12 become more and more questionable as they sweep from left to right. An examination of the adaptive-filter output trace is useful in estimating the attenuation of the signal proper at higher convergence rates. Figure IV-13 exhibits the beamsteer and adaptive filter output for the critical weak event when the convergence rate is 25% and the filter set is being updated. It appears that the signal alone has been knocked down by an amount consistent with the stated values in Table IV-3. Consequently, more confidence can be placed in the results derived solely from power measurements: for this signal, at least, maximum signal-to-noise gain does indeed occur near a convergence rate of 0.5%. Figure IV-13 also illustrates an artifact created by the high convergence rate. Toward the end of the adaptive-filter output in the bottom trace, there is a noticeable disturbance which is missing in the beamsteer output and the adaptive-filter output

TABLE IV-6

ADAPTIVE FILTERING SIGNAL-TO-NOISE GAIN VERSUS CONVERGENCE RATE FOR A WEAK SIGNAL FROM 300^{o}-305^{o} (USING NOISE REDUCTION MEASUREMENTS FROM DAY 238 OF 1970)

Convergence Rate (% of maximum)	Broadband Signal-to-Noise Gain (dB)	Narrowband Signal-to-Noise Gain (dB)
0.2	1.123	1.195
0.3	1.255	1.346
0.4	1.266	1.367
0.5	1.231	1.338
0.7	1.163	1.281
1.0	1.038	1.168
1.5	0.885	1.207
2.0	0.777	0.927
3.0	0.665	0.816
5.0	0.573	0.729
7.0	0.513	0.682
10.0	0.343	0.553
15.0	0.042	0.327
20.0	-0.187	0.130
25.0	-0.284	0.037

TABLE IV-7

ADAPTIVE FILTERING SIGNAL-TO-NOISE GAIN VERSUS CONVERGENCE RATE FOR A WEAK SIGNAL FROM 300°-305° (USING NOISE REDUCTION MEASUREMENTS FROM DAY 203 OF 1971)

Convergence Rate (% of maximum)	Broadband Signal-to-Noise Gain (dB)	Narrowband Signal-to-Noise Gain (dB)
0.2	1.729	1.867
0.3	1.930	2.088
0.4	2.020	2.170
0.5	2.034	2.188
0.7	2.036	2.197
1.0	1.992	2.167
1.5	1.906	2.109
2.0	1.851	2.080
3.0	1.812	2.091
5.0	1.800	2.140
7.0	1.744	2.106
10.0	1.533	1.916
15.0	1.106	1.512
20.0	0.791	1.198
25.0	0.638	0.984

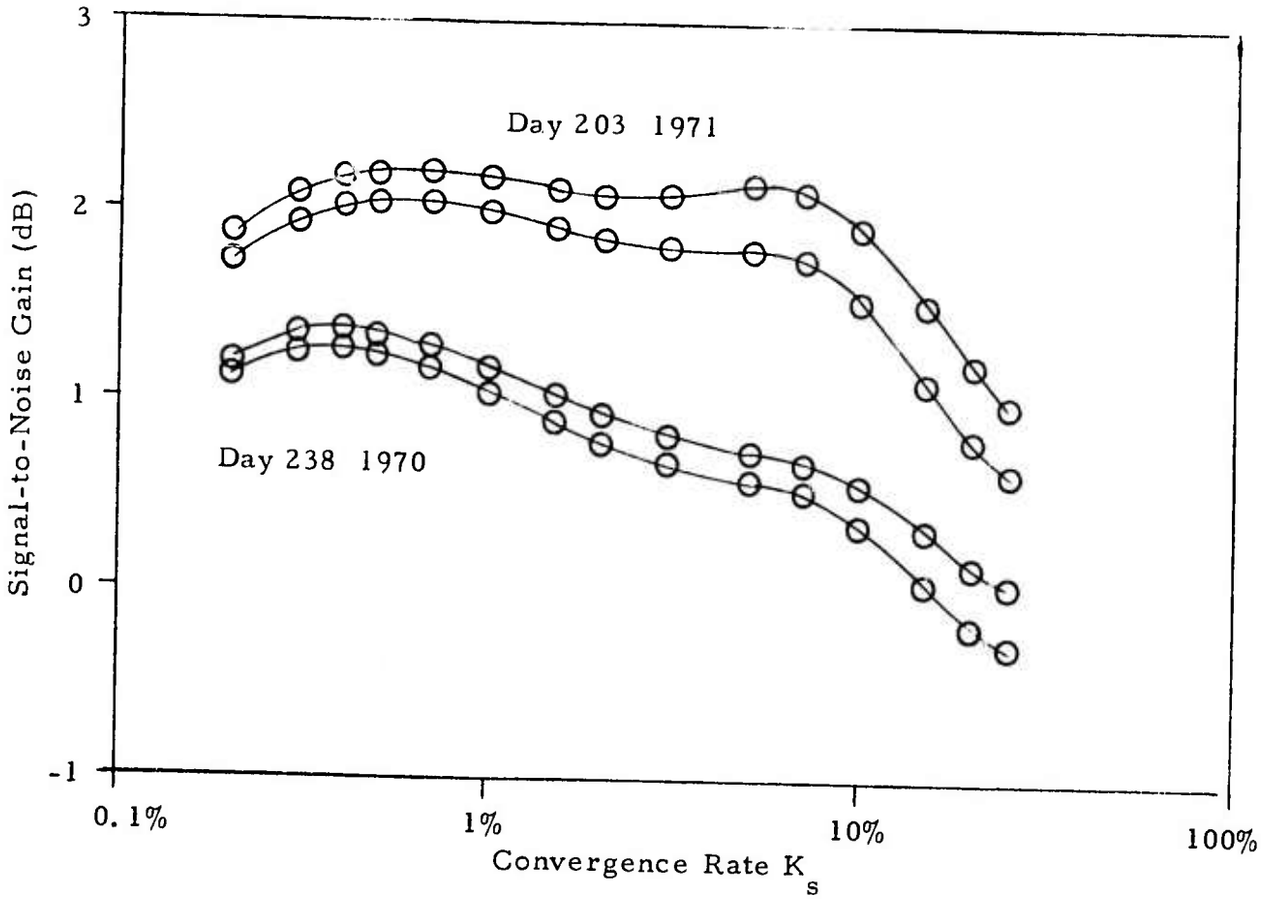


FIGURE IV-12

SIGNAL-TO-NOISE GAIN AS A FUNCTION OF CONVERGENCE RATE FOR A WEAK SIGNAL APPROXIMATELY 6 dB ABOVE NOISE LEVEL ON BEAMSTEER OUTPUT (STEER AZIMUTH 302.5°)

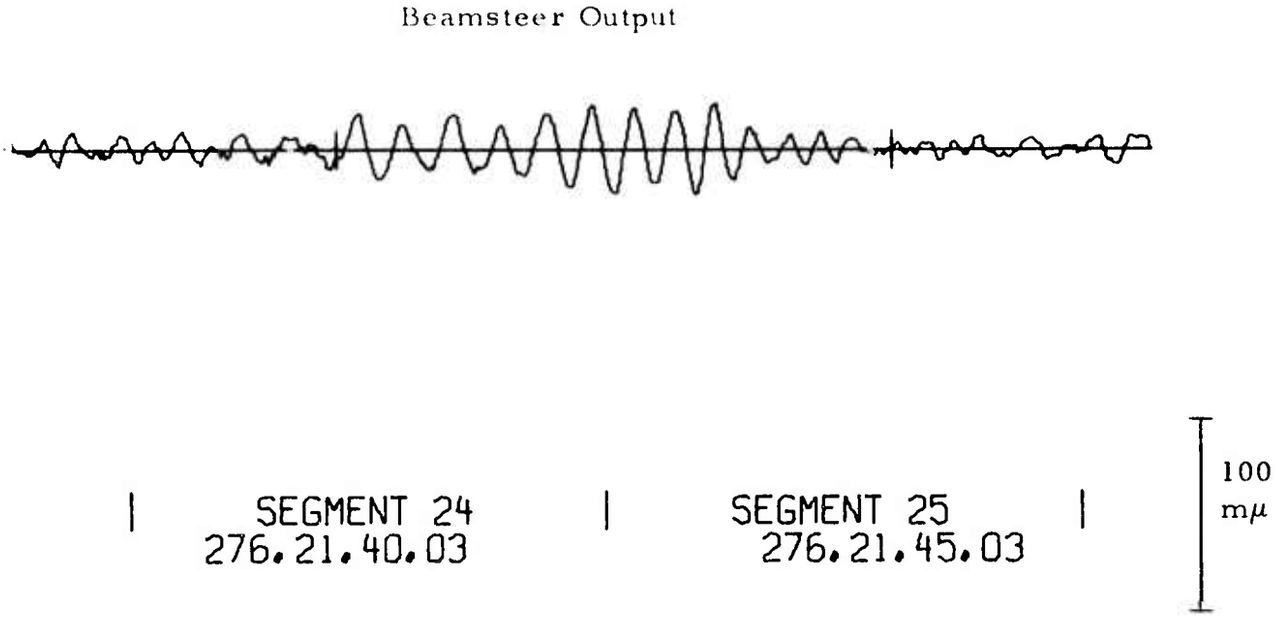


FIGURE IV-13

WEAK EVENT FROM 300°-305°
(FILTER ADAPTING, STEER DIRECTION 302.5°, K_s = 0.25)

at a convergence rate of 0.5% (see Figure IV-8). Apparently the adaptive-filtering process is attempting to use filtered noise in order to cancel the signal which has just died away.

In the remainder of this subsection, the two strong Kamchatka signals are treated in tandem to demonstrate the greater signal degradation and corresponding lower signal-to-noise ratio improvement for signals with increasingly higher signal-to-noise ratios. It is again worth emphasizing that first the beamsteer output is available for subsequent analysis when adaptive filtering spoils a strong signal and second, if necessary, the situation can be alleviated simply by freezing the adaptive filter set upon detection of a signal.

Tables IV-8 and IV-9 present the signal-to-noise ratio improvement using noise measurements from day 238 of 1970 for the earlier 18 dB signal and the later 24 dB signal, respectively. Figure IV-14 combines the results given in both tables. Corresponding results for day 203 of 1971 are contained in Tables IV-10 and IV-11 and Figure IV-15. On both days signal degradation is much worse for the later and stronger event. Positive improvement for the larger signal is achieved only at extremely low convergence rates. Maximum improvement occurs near a convergence rate of 0.2%. There broadband and narrowband signal-to-noise gain for the larger event are 0.136 dB and 0.208 dB, respectively, for day 238 of 1970, and 0.742 dB and 0.880 dB, respectively, for day 203 of 1971. The signal-to-noise ratio improvement values for the smaller Kamchatka signal are more in line with those of the 6-dB signal from 300° - 305°: the best results are near a convergence rate 0.5% of maximum. For both the strong Kamchatka signals, the worst improvement over beamsteering occurs at a 5% convergence rate (unlike the weaker event from 300° - 305°).

An interesting phenomenon happens as the convergence rate increases above 5%: the gain for the two Kamchatka earthquakes begins to rise again. Narrowband signal-to-noise ratio improvement actually exceeds that achieved

TABLE IV-8

ADAPTIVE FILTERING SIGNAL-TO-NOISE GAIN VERSUS CONVERGENCE RATE FOR A STRONG SIGNAL FROM KAMCHATKA (USING NOISE REDUCTION MEASUREMENTS FROM DAY 238 OF 1970)

Convergence Rate (% of maximum)	Broadband Signal-to-Noise Gain (dB)	Narrowband Signal-to-Noise Gain (dB)
0.2	0.830	0.888
0.3	0.943	1.029
0.4	0.965	1.064
0.5	0.946	1.072
0.7	0.869	1.021
1.0	0.704	0.888
1.5	0.443	0.663
2.0	0.209	0.456
3.0	-0.083	0.197
5.0	-0.288	0.046
7.0	-0.213	0.176
10.0	0.067	0.543
15.0	0.428	1.023
20.0	0.589	1.243
25.0	0.649	1.349

TABLE IV-9

ADAPTIVE FILTERING SIGNAL-TO-NOISE GAIN VERSUS CONVERGENCE RATE FOR A VERY STRONG SIGNAL FROM KAMCHATKA (USING NOISE REDUCTION MEASUREMENTS FROM DAY 238 OF 1970)

Convergence Rate (% of maximum)	Broadband Signal-to-Noise Gain (dB)	Narrowband Signal-to-Noise Gain (dB)
0.2	0.136	0.208
0.3	0.035	0.134
0.4	-0.134	-0.015
0.5	-0.338	-0.208
0.7	-0.758	-0.614
1.0	-1.377	-1.229
1.5	-2.264	-2.135
2.0	-2.948	-2.845
3.0	-3.732	-3.677
5.0	-4.044	-4.002
7.0	-3.856	-3.782
10.0	-3.405	-3.263
15.0	-2.918	-2.692
20.0	-2.709	-2.437
25.0	-2.605	-2.299

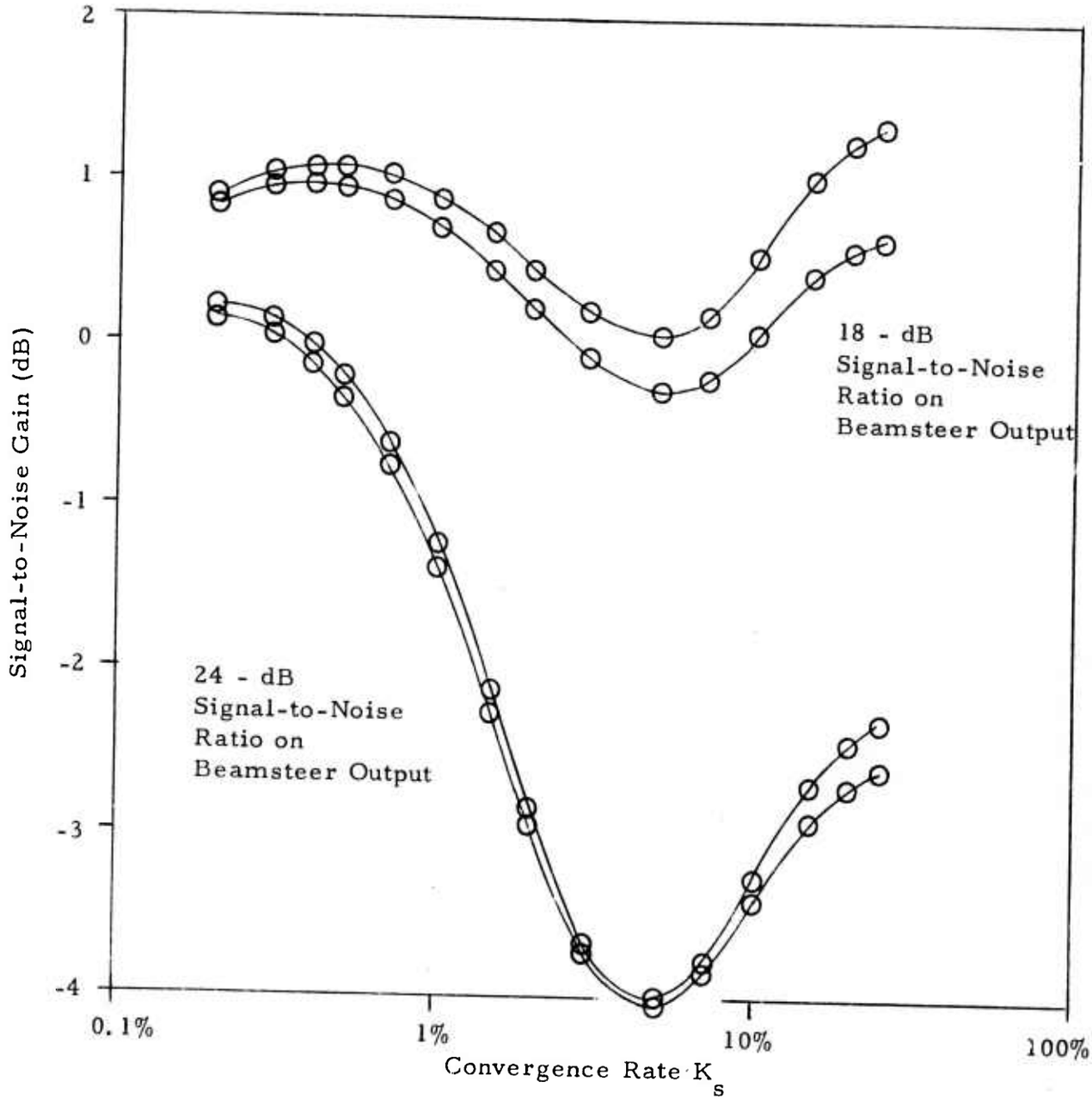


FIGURE IV-14

SIGNAL-TO-NOISE GAIN AS A FUNCTION OF CONVERGENCE RATE FOR TWO STRONG KAMCHATKA SIGNALS (USING NOISE REDUCTION MEASUREMENTS FROM DAY 238 OF 1970)

TABLE IV-10

ADAPTIVE FILTERING SIGNAL-TO-NOISE GAIN VERSUS CONVERGENCE RATE FOR A STRONG SIGNAL FROM KAMCHATKA (USING NOISE REDUCTION MEASUREMENTS FROM DAY 203 OF 1971)

Convergence Rate (% of maximum)	Broadband Signal-to-Noise Gain (dB)	Narrowband Signal-to-Noise Gain (dB)
0.2	1.436	1.560
0.3	1.628	1.771
0.4	1.719	1.877
0.5	1.749	1.922
0.7	1.742	1.937
1.0	1.658	1.887
1.5	1.464	1.745
2.0	1.283	1.609
3.0	1.064	1.472
5.0	0.939	1.457
7.0	1.018	1.600
10.0	1.257	1.906
15.0	1.492	2.208
20.0	1.561	2.311
25.0	1.571	2.333

TABLE IV-11

ADAPTIVE FILTERING SIGNAL-TO-NOISE GAIN VERSUS CONVERGENCE RATE FOR A VERY STRONG SIGNAL FROM KAMCHATKA (USING NOISE REDUCTION MEASUREMENTS FROM DAY 203 OF 1971)

Convergence Rate (% of maximum)	Broadband Signal-to-Noise Gain (dB)	Narrowband Signal-to-Noise Gain (dB)
0.2	0.742	0.880
0.3	0.720	0.876
0.4	0.620	0.788
0.5	0.465	0.642
0.7	0.115	0.302
1.0	-0.423	-0.320
1.5	-1.243	-1.053
2.0	-1.874	-1.692
3.0	-2.585	-2.402
5.0	-2.817	-2.591
7.0	-2.625	-2.358
10.0	-2.215	-1.900
15.0	-1.854	-1.507
20.0	-1.731	-1.369
25.0	-1.683	-1.315

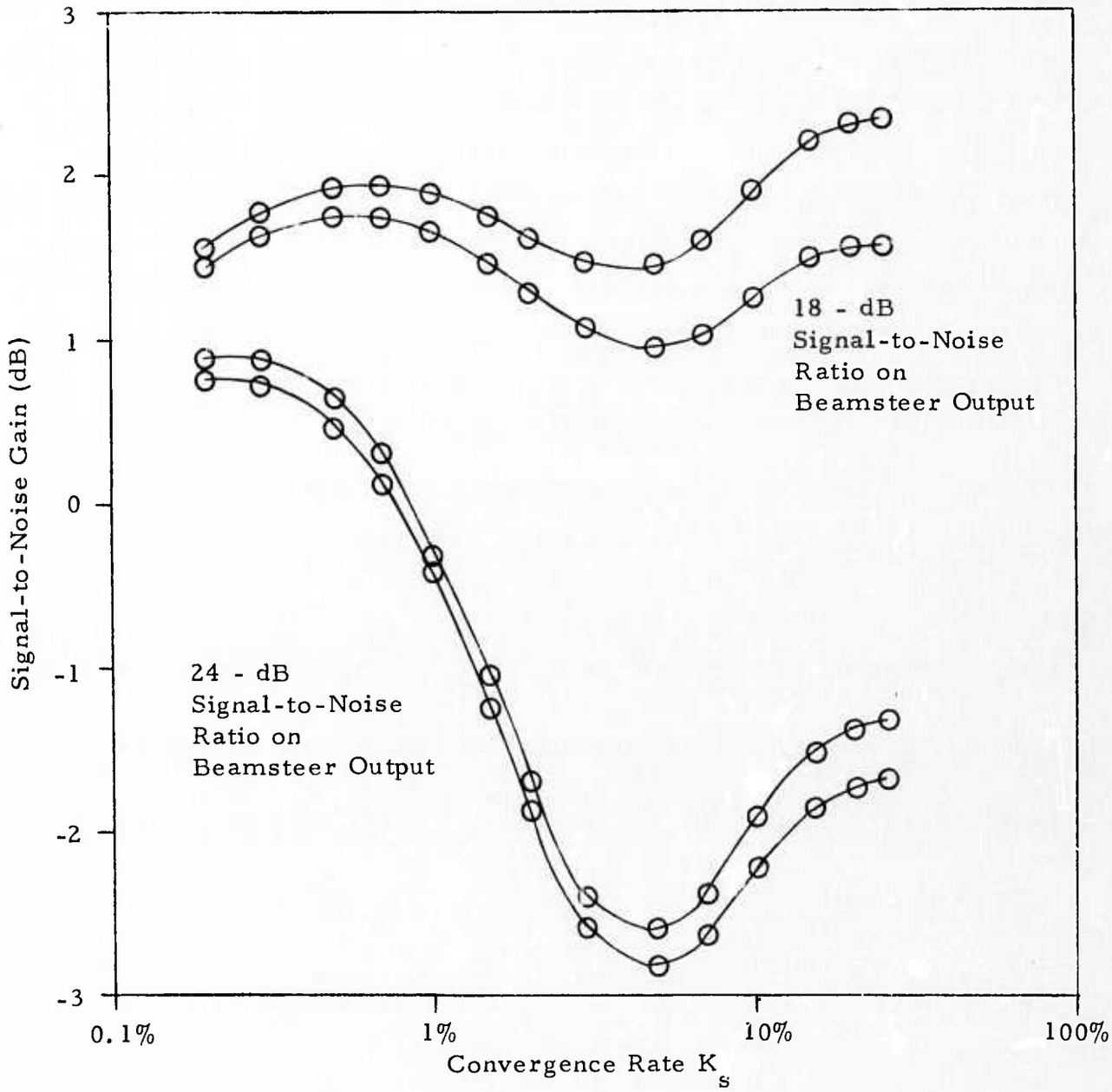


FIGURE IV-15

SIGNAL-TO-NOISE GAIN AS A FUNCTION OF CONVERGENCE RATE FOR TWO STRONG KAMCHATKA SIGNALS (USING NOISE REDUCTION MEASUREMENTS FROM DAY 203 OF 1971)

near 0.5% in the case of the 18-dB Kamchatka signal. The rise in gain is due to a flattening of signal degradation at convergence rates below those where maximum noise reduction is obtained (see Figures IV-9 and IV-10). This leveling off contrasts with signal degradation which is still climbing at a 25% convergence rate in the case of the weaker 6-dB signal from 300^{o}-305^{o} (see Figure IV-7). At any rate, there does not seem to be any way to exploit the extra noise reduction derived from transient correlation between successive data vectors. Such noise reduction would cease abruptly, anyhow, if the filter were frozen and the adaptive-filtering process could no longer track short-term fluctuations in the data. In addition, severe signal distortion is a compelling reason for avoiding the higher convergence rates.

The next section of this report discusses the effects of stopping the filter update in an attempt to eliminate the severe degradation of easily detected signals.

SECTION V

EFFECT OF FILTER FREEZE ON SIGNAL-TO-NOISE RATIO IMPROVEMENT

A. GENERAL DISCUSSION

The maximum-likelihood adaptive-filtering equations are designed to preserve a signal by constraining the filter set to have unity response in the signal direction while optimally reducing the total filter output power subject to the unity-response constraint. When no signal is present, the adaptive filter set optimally reduces the filtered noise power subject to the constraint conditions. There are no inherent design problems unless a signal is present. If a signal is present, however, the adaptive filter set will attempt to minimize the total output power by manipulating the filtered noise output so as to cancel as much of the filtered signal as possible while still observing the constraint conditions. The constraint conditions are a significant but not insurmountable barrier to signal cancellation. In the particular form of the update algorithm used in this study, furthermore, the step size of the filter update vector increases without limit as the time-shifted input data channels become more and more alike: the squared magnitude of the filter change vector was shown in Subsection E of Section III to be

$$|A^{new} - A^{old}|^2 = \frac{4K_s^2 y^2(t)}{|\bar{X}-X|^2} ,$$

where $y(t)$ is the adaptive filter output; as the data vector X approaches the beamsteer vector $\bar{X}$ (see page III-1 for the definitions), the step size $|A^{new} - A^{old}|$ approaches infinity. (In practice, the step size is limited by

contaminating noise, differences in instrument response, and signal propagation across the array not consistent with the signal model.) To prevent signal cancellation and filter instability, the adaptive filter set is frozen when a signal is detected. After the signal has died out, the filter set is again allowed to adapt with the aim of reducing the output power of a changed noise field.

A simple power detector was originally tried as a detection algorithm. A running power average was computed for the adaptive-filter output beam. Each point of the beam was checked to see whether its square exceeded some arbitrary multiple of the running power average for the beam. There were two serious problems with this method of detection:

- The running power average tended to increase with time, even when it was not updated after a signal detection. The time required to fall from a peak level was greater than the time required to climb to that peak level. Signals just below the detection threshold then drove the running power average still higher (and thereby raised the detection threshold).
- Noise fluctuations and glitches produced frequent false alarms when the threshold was set low enough to detect signals clearly visible on the beam output trace.

A detection algorithm was chosen specifically to eliminate excessively large jumps in the adaptive filter vector. For a fixed convergence rate K_s, the step size is proportional to $|y(t)| / |\bar{X}-X|$. When this ratio is large, it is likely that a signal is present. To eliminate the dependence of this ratio on the filter weights, the adaptive filter output $y(t) = X^T A$ is replaced by the RMS beamsteer output over a 31-point, 31-second gate. The final form of the detection ratio is

$$\frac{\bar{X}^T\bar{X}}{(\bar{X}-X)^T(\bar{X}-X)} = \frac{\sum_{i=1}^{M}\sum_{j=-N}^{N}\bar{x}^2(t-j)}{\sum_{i=1}^{M}\sum_{j=-N}^{N}\left[\bar{x}(t-j) - x_i(t-j)\right]^2}$$

where M is the number of channels, 2N+1 is the total number of filter weights per channel, $\bar{x}(t-j)$ is the beamsteer output at time $t-j\Delta t$, and $x_i(t-j)$ is the prefiltered vertical-component output for site i at time $t-j\Delta t$. This ratio is a measure of the similarity between the input channels and the beamsteer output. As it turns out, the detection ratio is a scaled version of the Fisher detector $\left[(M-1)\bar{X}^T\bar{X}\right] / \left[(\bar{X}-X)^T(\bar{X}-X)\right]$. The Fisher detector has been studied extensively (Melton and Bailey, 1957; Booker, 1965; Edwards, Benno and Creasey, 1967; Shumway and Husted, 1970; Blandford, 1970; Shumway, 1971; Wirth, 1971; Smart and Flinn, 1971; Wirth, Blandford and Shumway, 1971; Shumway, 1972; Blandford, 1972; Smart, 1972). In the results which follow, the threshold value for the detection ratio is 4 (sufficient to detect signals 12 dB above the noise on the beamsteer output for six sites at ALPA). Once the detection ratio exceeds 4, the filter set is frozen until 120 successive detection ratios (or two minutes of data) fall below the threshold.

Subsection B shows what happens to the two large Kamchatka events from the previous section when the adaptive filter is frozen after a signal detection. In addition, it illustrates the capability of adaptive multichannel filtering to suppress off-azimuth events better than the standard time-shift-and-sum beamforming method. Subsection C tackles the difficult problem of estimating the loss in noise reduction during the time period when the adaptive filter is prevented from updating. Subsection D describes an alternate adaptive

algorithm which appears to have definite advantages over the algorithm actually used. The principal advantage of the suggested algorithm is greatly reduced signal degradation.

B. EFFECT OF FILTER FREEZE ON SIGNALS

A magnitude 5.9 earthquake from New Guinea on day 7 of 1972 was selected to test the filter freeze procedure and observe its effects on the signal. The PDE bulletin gives its origin time as 06:25:48.4 and its location as (2.1S, 139.0E). The azimuth from ALPA is 253°. Both the beamsteer and adaptive-filter beams were aimed in this direction to pass energy of 3.5km/sec velocity. To prevent clipping of this signal, which was more than 42 dB above the noise on the beamsteer output, the data was scaled by a factor of only 4 instead of 16. The adaptive-filter package was run from 0430 to 0826 at a convergence rate of 0.5% using sites 8, 9, 12, 15, 16, and 17 as input to the adaptive-filter beam.

The similarity-detection algorithm had no difficulty in spotting this event: the adaptive filter set was frozen several times during the Rayleigh wave arrivals from this event. The main Rayleigh wave arrived at about 0708 and is shown in Figure V-1. Tick marks at zero level in both beams are separated by 5 minutes and indicate the times given midway between traces. The time-shift-and-sum output is at the top, the adaptive-filter output at the bottom. The Rayleigh-wave arrival for which both beams are stored is virtually intact. Signal degradation measured over a 512-point gate starting at the beginning of segment 39 in the figure was just under 1.3 dB both broadband and in the frequency band associated with periods between 43 and 15 seconds. This event is so strong that the P-wave arrival and other seismic phases ahead of the Rayleigh wave are visible on the beam outputs. Since no appreciable degradation occurs for other events during the freeze period when the Rayleigh wave is the only visible phase, it is likely that the adaptive-filtering process is rejecting some of the Rayleigh-wave arrival on an azimuthal

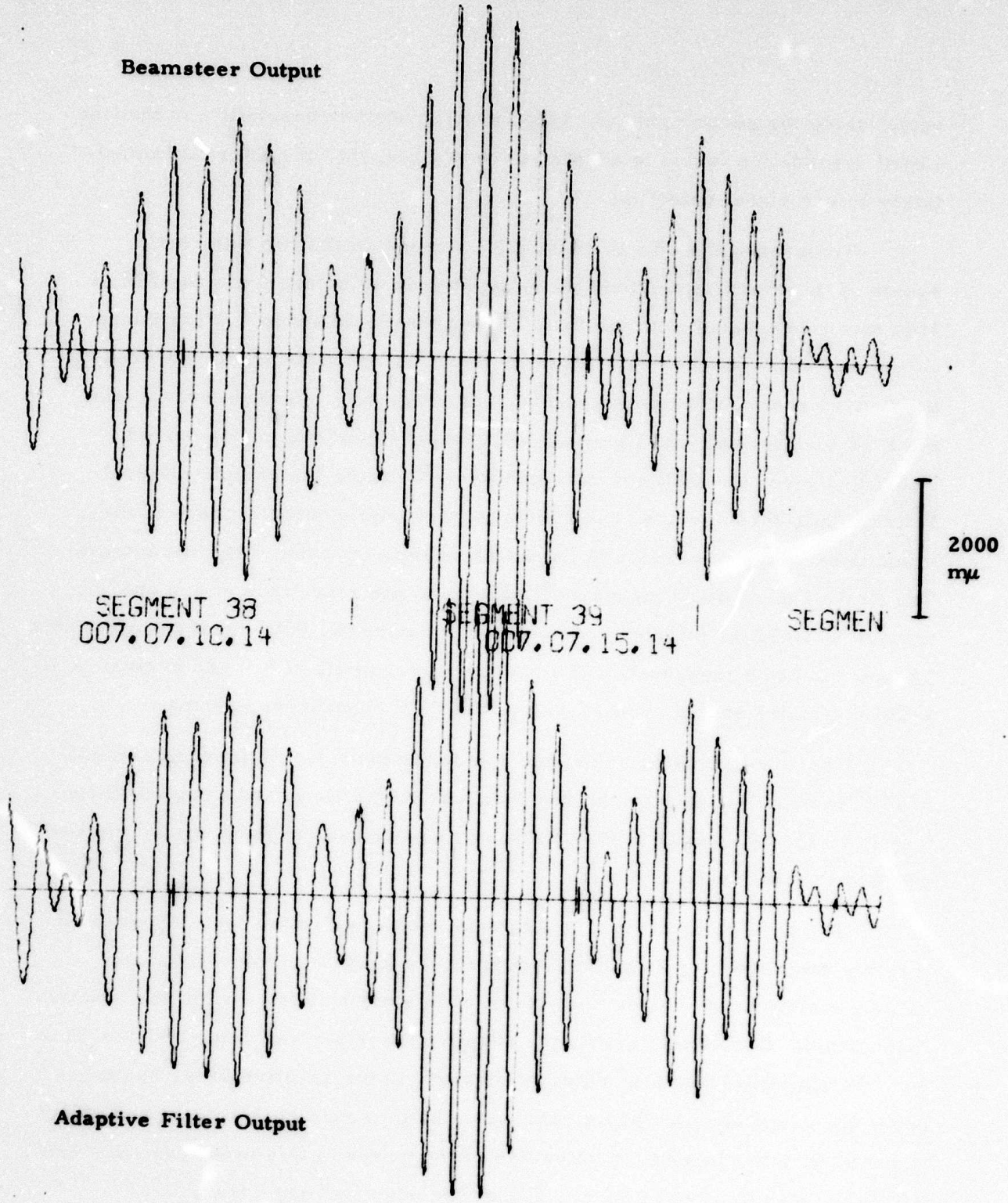


FIGURE V-1

MAIN RAYLEIGH WAVE ARRIVAL FROM 5.9 NEW GUINEA EVENT (JANUARY 7, 1972)

basis (using the earlier phases of the event). Another possibility is that the signal degradation is due to adaptation over the short time interval immediately before signal detection.

To observe the effects of the filter freeze on the two Kamchatka events of the previous section, the same sample as before from day 276 of 1971 was reprocessed with the filter freeze threshold set to 4. The convergence rate was specified as 0.5% of maximum in the runs which produced the figures shown in the remainder of this subsection. Several events from the PDE bulletin and LASA bulletin had arrival times within this period. Table V-1 gives the PDE events which arrived during the sample interval. In addition to these events, two events located within three degrees of the Panama event were included in the LASA bulletin covering this time interval. The first of these had a measured LASA bodywave magnitude of 3.6 and arrived at LASA 47 minutes and 34 seconds before the PDE event from Panama. The second had a measured LASA bodywave magnitude of 3.9 and arrived at LASA 6 minutes and 38 seconds before the PDE event from Panama.

The adaptive filter beam was steered to pass events from an azimuth of 273° in order to pick up the two events from the Kamchatka region. Sites 8, 9, 12, 15, 16, and 17 of the ALPA array were used to form the beamsteer output and adaptive-filter beam.

The first clear arrival at ALPA is shown in Figure V-2. Apparently it is the magnitude 3.6 Panama event from the LASA bulletin. Although clearly visible in the beamsteer output, it is almost obliterated in the adaptive-filter beam. During this event, the adaptive filter set was being updated. This off-azimuth signal was attenuated so strongly in the adaptive-filter beam because the adaptive-filter beam reacted to the presence of this event and nulled it out: a steadily increasing reduction of the power in this event over the first few cycles can be seen in Figure V-2 on the adaptive-filter trace.

TABLE V-1

PDE EVENTS ARRIVING AT ALPA BETWEEN 2000 AND 2357
ON OCTOBER 3, 1971

Origin Time (Hr-Min-Sec)	Lat (Deg)	Long (Deg)	Region	m_b	Azimuth (Deg)
20:42:46.7	3.5N	82.9W	South of Panama	4.7	113
20:54:48.5	55.7N	162.1E	Near East Coast of Kamchatka	4.5	273
21:54:12.9	55.8N	162.2E	Near East Coast of Kamchatka	4.9	273
22:34:54.8	4.2S	152.7E	New Britain Region	4.8	242
23:17:26.4	51.8N	173.3W	Andreanof Islands, Aleutian Is.	4.3	237

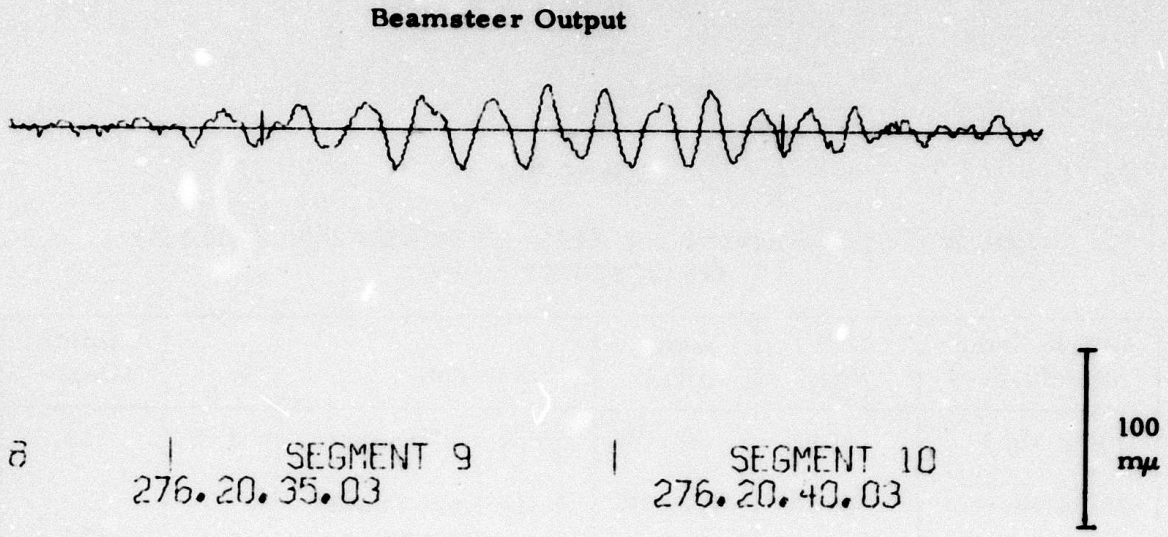


FIGURE V-2

EVENT PROBABLY FROM SOUTH OF PANAMA (LASA m_b = 3.6)

The first of the two on-azimuth events from Kamchatka is plotted in Figure V-3. The filter-freeze algorithm prevented the adaptive filter set from being updated during this event. As a result, signal degradation for this event was less than 0.1 dB. There is almost no perceptible difference between the beamsteer and adaptive-filter traces in Figure V-3.

The magnitude 3.9 Panama event from the LASA bulletin did not appear on either beam.

On the other hand, the magnitude 4.7 Panama event from the PDE bulletin was clearly evident on the beamsteer output. This event is pictured in Figure V-4. As in the case of the event 47 minutes earlier, it is significantly weaker on the adaptive-filter beam. The adaptive filter set was updated throughout this event.

Figure V-5 displays the weak event from 300°-305° used in the previous section for signal degradation measurements. In this case, however, the 273° steer direction transforms this event into an off-azimuth event. Attenuation is 0.56 dB broadband and 0.57 dB narrowband when the detection of the first Kamchatka event in Figure V-3 triggers filter-update suppression over that event. Curiously enough, the attenuation for this weak event from 300°-305° is -0.35 dB broadband and -0.36 dB narrowband when the adaptive filter is permitted to vary during the preceding Kamchatka event. That is to say, when the only difference in processing is to permit or to suppress the adaptive-filter update algorithm during an on-azimuth event 30 minutes earlier, the resultant attenuation varies by almost 0.9 dB. In both cases, the steer direction is 273°, the convergence rate is 0.5%, the same sites are utilized, and the filter set is adapting for at least 20 minutes after the earlier Kamchatka event has faded away. Figure V-6 presents the same weak off-azimuth event after allowing the filter set to vary over the previous event from Kamchatka. This figure does indeed appear to contain a cleaner estimate of the signal. A possible explanation is that the adaptive-filter beam pattern exceeds

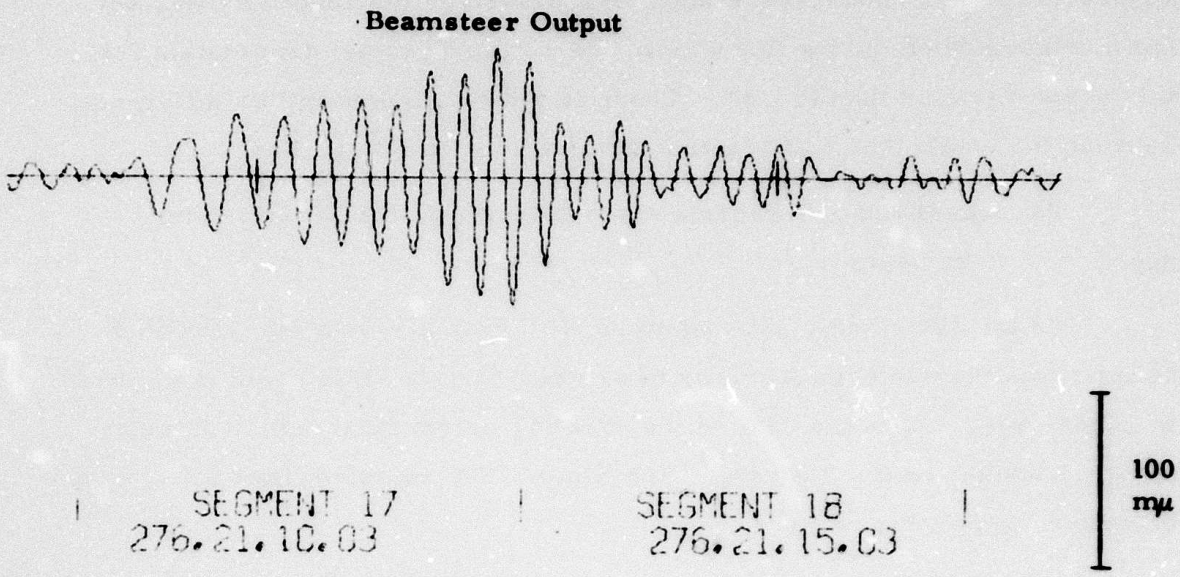


FIGURE V-3

MAGNITUDE 4.5 EVENT FROM KAMCHATKA (FILTER FROZEN, STEER DIRECTION 273°, K_s = 0.005)

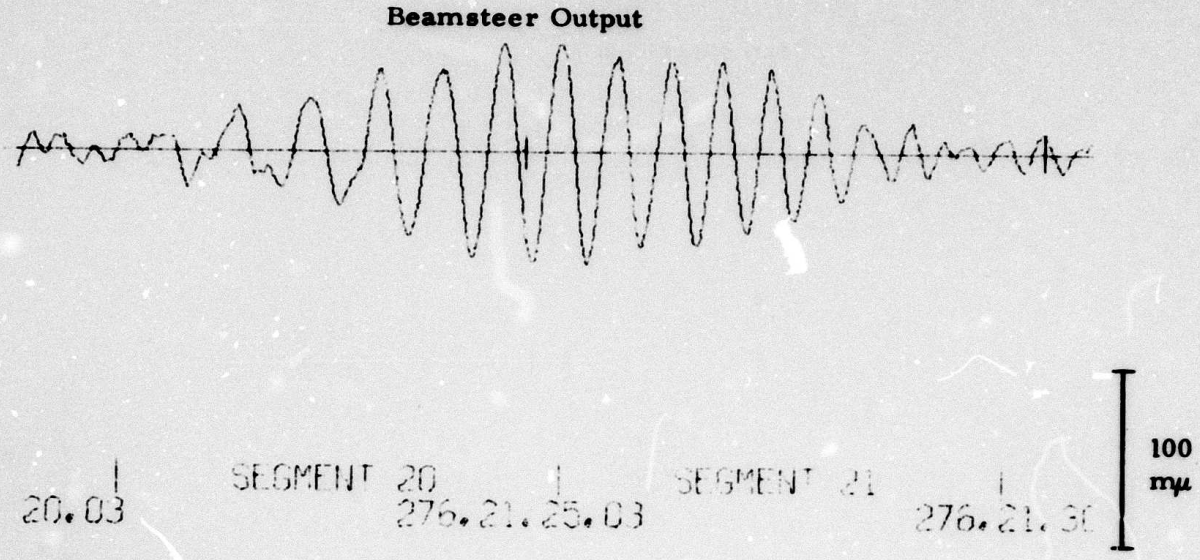


FIGURE V-4

MAGNITUDE 4.7 EVENT FROM SOUTH OF PANAMA

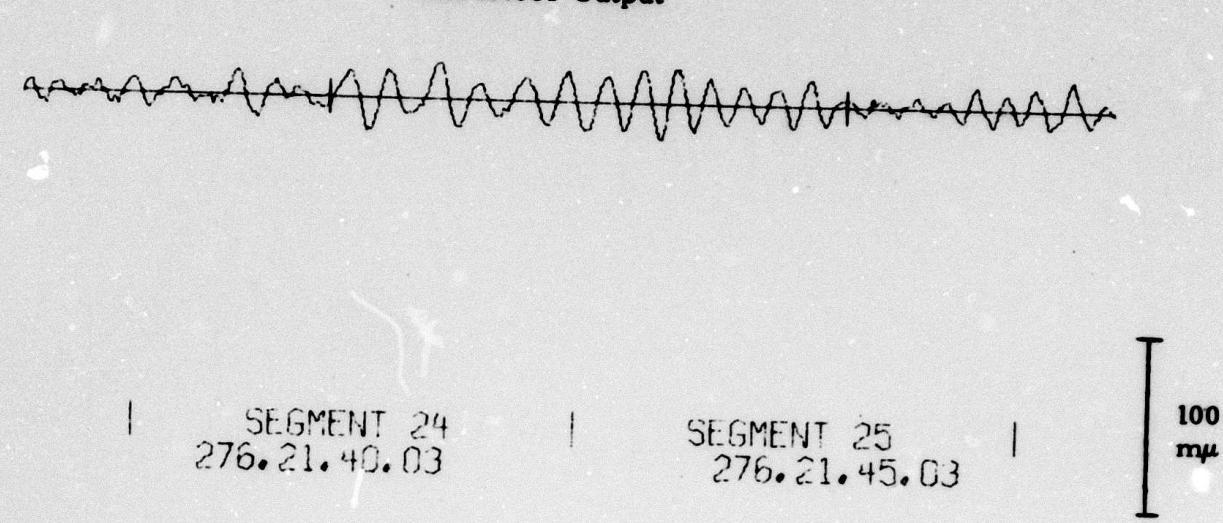


FIGURE V-5

WEAK EVENT FROM 300°-305°
(FILTER ADAPTING, STEER DIRECTION 273°, $K_g = 0.005$,
FILTER FROZEN DURING EARLIER KAMCHATKA EVENT)

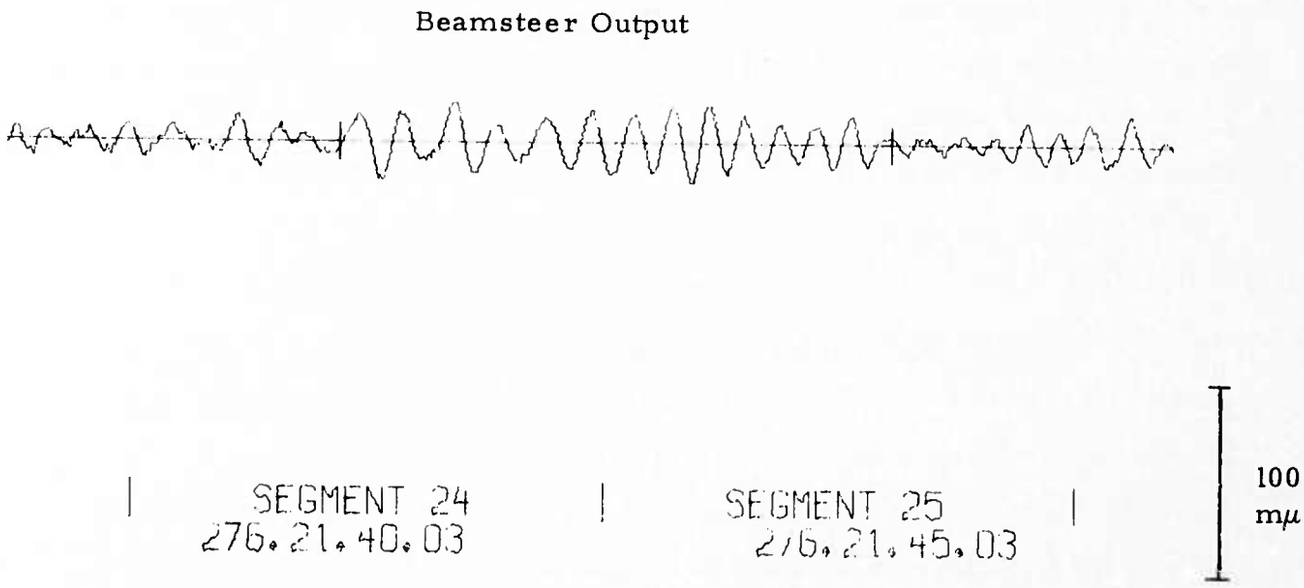


FIGURE V-6

WEAK EVENT FROM 300°-305°
(FILTER ADAPTING, STEER DIRECTION 273°, $K_s = 0.005$,
FILTER ADAPTING DURING EARLIER KAMCHATKA EVENT)

a value of 0 dB at 300°-305° azimuths at the time of arrival of the off-azimuth event in Figures V-5 and V-6. Array response patterns with maxima away from the steer direction are often generated by multichannel filtering, especially when no significant energy is coming from the direction corresponding to the beam-pattern maximum. If this interpretation is correct, it is still a moot question whether the deflection of the main lobe to azimuths north of 273° had already taken place by the end of the Kamchatka event or whether the characteristics of the intervening Panama event contributed to the end result.

The second Kamchatka event from the PDE bulletin is presented in Figure V-7. The filter set was frozen during this event. This figure should be compared with Figure IV-11, where the filter is adapting. Signal degradation for this on-azimuth event is almost exactly 0 dB. As in the case of the earlier Kamchatka event, there is almost no difference between the two beam outputs.

No definite detection could be made for the magnitude 4.8 event from the New Britain region.

The final event from the PDE bulletin, a magnitude 4.3 event from the Andreanof Islands (Figure V-8), was detected on both the beamsteer and adaptive-filter beams. However, the similarity between the beamsteer output and the input channels was too low to freeze the adaptive filter set. This earthquake illustrates the effect of adaptive filtering on a strong off-azimuth signal reasonably close (36°) to the steer direction. Despite the loss of one input channel and the resultant redistribution of filter weights just before the principal burst of energy on both traces, attenuation was more than 6 dB relative to the beamsteer output. The marked attenuation demonstrates the sometimes-forgotten capability of multichannel filtering to narrow the width of the main lobe in the array beam pattern: the array becomes a superdirective antenna.

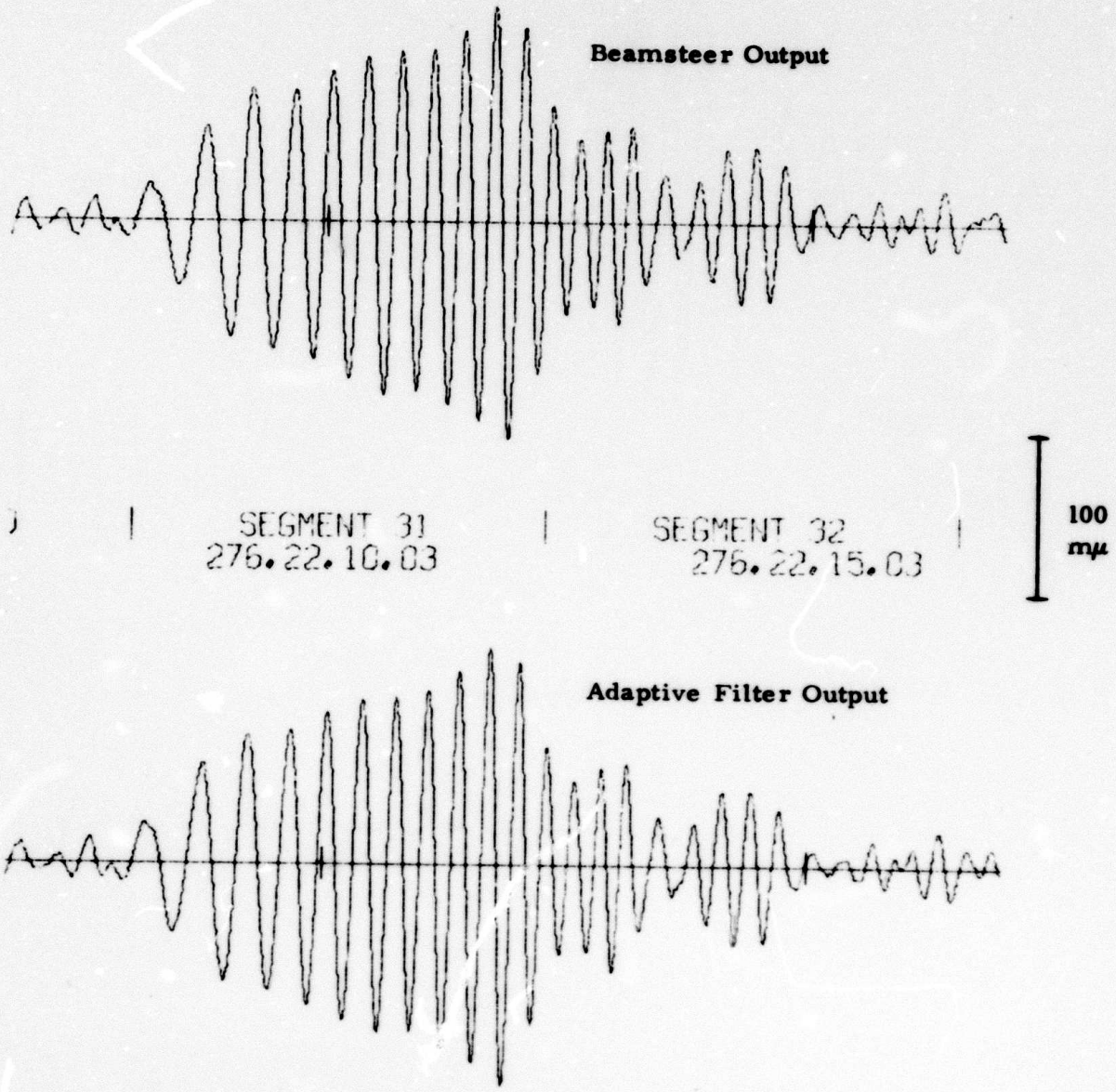


FIGURE V-7

MAGNITUDE 4.9 EVENT FROM KAMCHATKA (FILTER FROZEN, STEER DIRECTION 273°, $K_s = 0.005$)

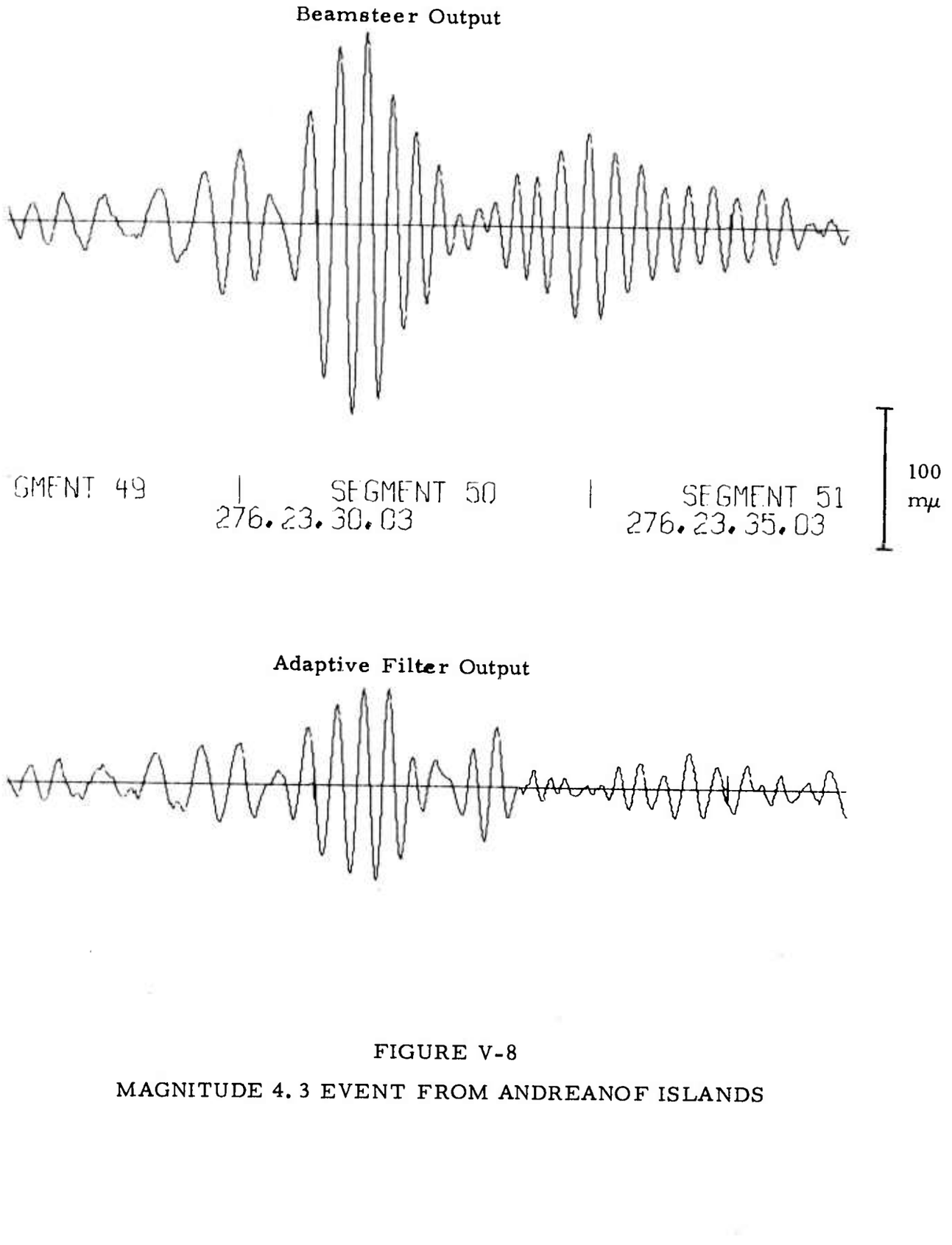


FIGURE V-8

MAGNITUDE 4.3 EVENT FROM ANDREANOF ISLANDS

The following conclusions can be drawn from this four-hour data sample:

- On-azimuth signals at ALPA are not significantly attenuated by adaptive filtering at convergence rates near optimum for suppression of background noise. Attenuation was 0.4 dB for the weak signal from 300°-305° (see Table IV-3) with the filter adapting and steered toward 302.5°. To stop degradation of stronger signals, the filter must be frozen.

- Off-azimuth events are strongly suppressed in the adaptive-filter beam when the filter set is not frozen. Some off-azimuth signals are virtually annihilated. The more powerful the off-azimuth event, the more it is smothered.

- Greater directional resolution at ALPA and other similar long-period arrays is achievable through the beam-narrowing capability of multichannel filtering. This fact is extremely important if long-period arrays of this type are to be used for signal detection and location or for separation of multiple events. With time-varying adaptive filters, off-azimuth events can be nulled out in proportion to their signal-to-noise ratio with possible complications if two events overlap in time. With fixed non-varying multichannel filters, superdirectivity can be preserved in all circumstances, but then the ability to quell specific bursts of off-azimuth energy in an on-line processing mode is impaired.

C. EFFECT OF FILTER FREEZE ON NOISE REDUCTION

In Section IV, signal-to-noise ratio improvement was calculated by subtracting signal degradation from noise reduction when the adaptive filter set is permitted to update each time a new sample of data is available, even

when an on-azimuth signal is detected. In this section, the method of determining signal-to-noise gain again involves the same two quantities, but in this case they must be measured so as to reflect the fact that the filter update has ceased. For signal degradation, the required procedure is relatively simple and was performed in Subsection B: only in the instance of the 42-dB signal on day 7 of 1972 was the degradation significant (1.3 dB); the 18-dB and 24-dB signals on day 276 of 1971 were attenuated less than 0.1 dB. The measurement of the loss in noise reduction, however, poses some problems. As the elapsed time increases from the point of the filter freeze, the general trend is one of even greater degradation in comparison with an adapting filter set. There are periods of time, however, when the frozen filter set is better able to reduce the noise power than the adapting filter set. The superior performance of the frozen filter set is due to temporary instances where the characteristics of the noise field revert to a condition more typical of the period before the filter freeze. An attempt is made to put bounds on the trend of the drop in noise reduction as a function of time elapsed since the filter freeze: this attempt is a subjective interpretation of the rapidly fluctuating measurements of the drop in noise reduction.

To determine the effect on noise reduction of freezing an adaptive filter set when an on-azimuth signal is detected, a four-hour noise sample from day 232 of 1970 was selected for processing. The time period covered is from 0344 to 0741. No events are reported by the PDE bulletin between 1219 on August 19 and 0834 on August 20 (day 232). A 360° Fisher-detector scan of the four-hour sample at 24° azimuthal increments suggests two possible signal arrivals, one at 0406 from a 252° azimuth and another at 0506 from an azimuth between 12° and 36°. In the first case, a faint dispersed wavetrain is visible between 0406 and 0416 on the 270° vertical-component beam. A trained analyst would probably call it a signal. The second possible arrival at 0506 is slightly weaker than the first on the Fisher-detector scan. The

only computed beam at 270° is too far away from the arrival direction to decide whether a signal is present. These two possible signals should have only minuscule effects on noise-reduction results during the critical final hour of the data sample.

In processing this sample, sites 3, 4, 5, 6, 8, and 9 are used to form vertical-component input channels for the beamsteer output and adaptive-filter beam. Both beams are steered to pass energy arriving from an azimuth of 270° at a velocity of 3.5 km/sec. The adaptive-filter convergence rate is set at 0.5% of maximum. The noise sample is processed twice. In both cases, the filter is permitted to adapt for the first three hours. In the first computer run, it is allowed to adapt for the remaining hour. In the second run, it is frozen for the remainder of the run in order to deduce the loss in noise reduction as a function of the elapsed time since the filter freeze. In both runs, time is divided into 55 equal 256-second segments. Segments 40-55 correspond to the period in which the adaptive-filtering update process is suppressed in the second computer run. Table V-2 gives the broadband noise reduction when the filter is adapting and when it is frozen, together with the difference between the two modes of operation. These figures are given for each of the 16 segments, for 8 two-segment intervals, 4 four-segment intervals, 2 eight-segment intervals, and the entire 16-segment period from 0629 to 0737. Figure V-9 plots the measured loss in noise reduction as a function of the time since the filter was frozen. Each point in the rightmost column of Table V-2 is located at the middle of the time interval it designates. Results are quite variable, so much so that the loss in noise reduction does not increase monotonically until eight segments are grouped together. The overall trend, of course, is toward ever greater loss of noise reduction. Figure V-10 attempts to portray the apparent trend of the drop in noise reduction due to freezing the adaptive filter set. It is assumed that the loss in noise reduction climbs monotonically from zero, starting at the point in time where the

TABLE V-2

BROADBAND NOISE REDUCTION AS A FUNCTION OF TIME FOR AN ADAPTING FILTER SET AND A FROZEN FILTER SET (USING A DATA SAMPLE FROM DAY 232 OF 1970)
(PAGE 1 OF 2)

Segment	Broadband Noise Reduction (dB)		Loss (dB)
	Filter Adapting	Filter Frozen	
40	2.124	1.811	0.313
41	1.419	1.262	0.157
42	1.191	0.703	0.488
43	1.803	1.150	0.653
44	1.348	1.088	0.260
45	3.305	2.931	0.374
46	0.688	-0.040	0.728
47	3.409	2.625	0.784
48	1.702	1.564	0.138
49	3.249	2.526	0.723
50	0.063	-0.710	0.773
51	1.227	1.736	-0.509
52	0.592	-0.202	0.794
53	2.649	1.742	0.907
54	0.031	-0.906	0.937
55	1.971	1.217	0.754

TABLE V-2

BROADBAND NOISE REDUCTION AS A FUNCTION OF TIME FOR AN ADAPTING FILTER SET AND A FROZEN FILTER SET (USING A DATA SAMPLE FROM DAY 232 OF 1970)
(PAGE 2 OF 2)

Segment	Broadband Noise Reduction (dB)		Loss dB
	Filter Adapting	Filter Frozen	
40-41	1.731	1.507	0.224
42-43	1.422	0.874	0.548
44-45	2.224	1.919	0.305
46-47	1.826	1.079	0.747
48-49	2.552	2.109	0.443
50-51	0.693	0.525	0.168
52-53	1.594	0.752	0.842
54-55	1.092	0.244	0.848
40-43	1.546	1.122	0.424
44-47	2.022	1.482	0.540
48-51	1.693	1.390	0.303
52-55	1.274	0.428	0.846
40-47	1.747	1.275	0.472
48-55	1.474	0.871	0.603
40-55	1.616	1.081	0.535

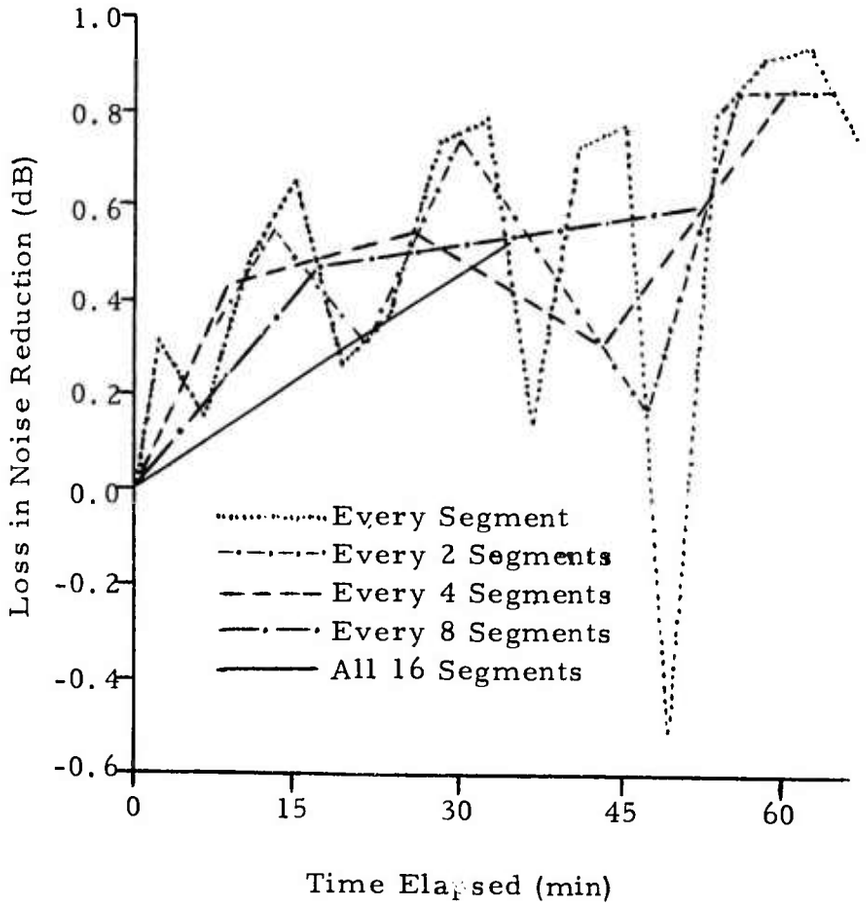


FIGURE V-9

MEASURED LOSS IN NOISE REDUCTION AS A FUNCTION OF TIME ELAPSED SINCE FILTER FREEZE (USING NOISE FROM DAY 232 OF 1970)

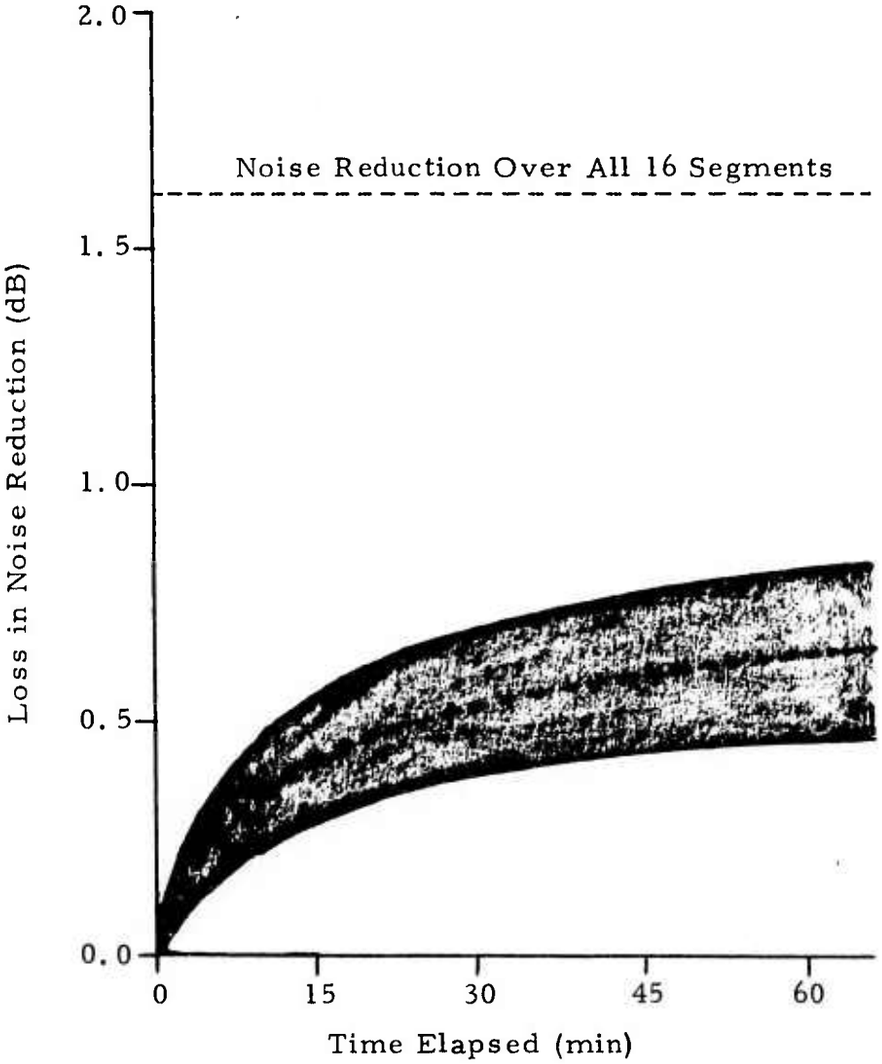


FIGURE V-10

APPARENT TREND OF NOISE REDUCTION LOSS AS A FUNCTION OF TIME ELAPSED SINCE FILTER FREEZE (USING NOISE FROM DAY 232 OF 1970)

update process ceases. The shaded area indicates the area between the likely upper and lower bounds for the trend. The dotted curve within the shaded area is the estimate of the trend. Shown at the top of the figure is the 1.6 dB noise reduction achieved by the adapting filter set over the interval 0629-0737.

To determine the signal-to-noise gain for a detected event, the signal degradation with the filter frozen is subtracted from the noise reduction to obtain the signal-to-noise gain at the time when the filter ceases to vary. For the day-276 Kamchatka signal 18 dB above noise level on the beamsteer output, broadband signal degradation was 0.087 dB. Therefore the 1.616 dB noise reduction over segments 40-55 of the sample from day 232 of 1970 at a convergence rate of 0.5% would yield a signal-to-noise ratio improvement of 1.529 dB at the time the filter is frozen. From this value, the drop in noise reduction due to freezing the filter must be subtracted. Figure V-11 shows the resultant trend of signal-to-noise gain as a function of time elapsed since the cessation of update using the results of Figure V-10. The gain which would have been achieved if the filter had not been frozen is 0.923 dB (1.616 dB - 0.693 dB) at a convergence rate 0.5% of maximum. Since the event in question lasted for less than 10 minutes, it is clearly desirable to suppress the filter update upon detection of this signal.

Even more dramatic are the results for the later Kamchatka event 24 dB above beamsteer noise level. Broadband signal degradation was -0.004 dB with the filter frozen, so that signal-to-noise ratio improvement falls off starting at a value of 1.620 dB. This figure contrasts with a signal-to-noise gain of -0.361 dB (1.616 dB - 1.977 dB) when the filter adapts at a 0.5% convergence rate.

In summary, when signals reach the signal-to-noise ratio required for detection, greater signal-to-noise gain is generated by the filter-freeze procedure. The advantages of preventing the filter update become more and more dramatic in the case of increasingly strong events.

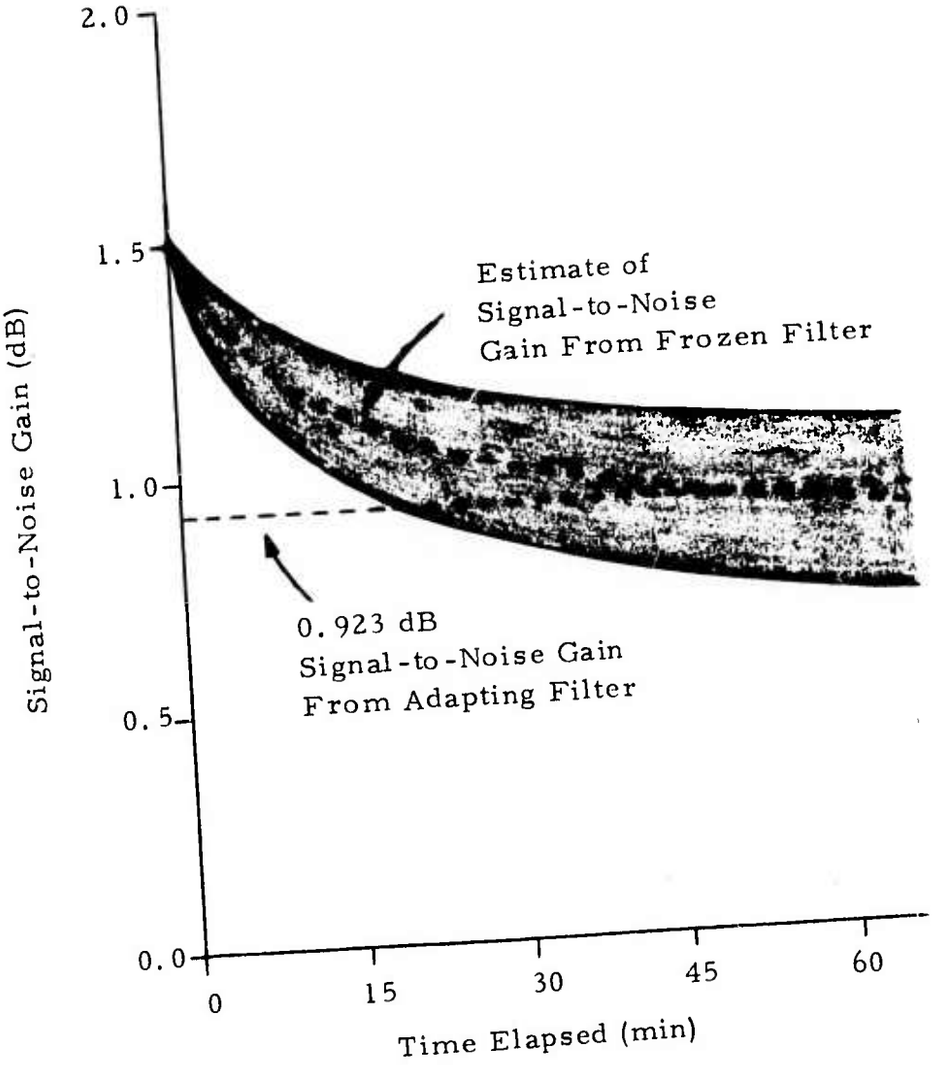


FIGURE V-11

TREND OF SIGNAL-TO-NOISE GAIN (WITH FILTER FROZEN) FOR KAMCHATKA EVENT 18 dB ABOVE BEAMSTEER NOISE LEVEL (USING NOISE FROM DAY 232 OF 1970)

D. AN ALTERNATE ADAPTIVE ALGORITHM

Part two of Subsection III-D showed that the adaptive algorithm

$$A^{new} = A^{old} + 2\mu X^T A^{old}(\bar{X}-X)$$

minimizes the mean square adaptive filter output $A^T E\left[XX^T\right]A$ subject to the constraint conditions $\sum_{i=1}^{M} a_i(j) = \delta_{oj}$ $(j=-N, \ldots, -1, 0, 1, \ldots, N)$ when the constant-valued convergence factor μ is chosen to be small enough. If such an algorithm had been used, bursts of power from off-azimuth seismic events would have been the dominant factor in the choice of the filter vectors A. Since seismic events cannot, in general, be expected to repeat themselves, adaptive filters designed on this basis would devote significant effort to the elimination of energy which had long since disappeared. The adaptive algorithm

$$A^{new} = A^{old} + \frac{2K_s X^T A^{old}(\bar{X}-X)}{(\bar{X}-X)^T(\bar{X}-X)}$$

used for this study attempts to minimize the quantity

$$A^T E\left[\frac{XX^T}{(\bar{X}-X)^T(\bar{X}-X)}\right] A$$

subject to the same maximum-likelihood constraints $\sum_{i=1}^{M} a_i(j) = \delta_{oj}$, where E denotes the expectation of each matrix element inside the brackets. To prevent the squared magnitude

$$\left|A^{new} - A^{old}\right|^2 = \frac{4K_s^2 y^2(t)}{(\bar{X}-X)^T(\bar{X}-X)}$$

of the filter update vector for this algorithm from becoming enormous when a strong signal traverses the array, the adaptive filter set must be frozen whenever a signal is detected. With this algorithm, each update reduces the power output for the old data vector by a factor $(1-2K_s)^2$ regardless of the fact that the data vector X may be consistent with a signal:

$$(A^T)^{new} XX^T A^{new} = (A^T)^{old} XX^T A^{old} (1-2K_s)^2 \quad .$$

The alternate algorithm

$$A^{new} = A^{old} + \frac{2K_s X^T A^{old} (\bar{X}-X)}{X^T X}$$

attempts to minimize the quantity

$$A^T E \left[\frac{XX^T}{X^T X} \right] A$$

subject to the same maximum-likelihood constraints. With this algorithm, the squared magnitude

$$\left| A^{new} - A^{old} \right|^2 = \frac{4K_s^2 y^2(t) (\bar{X}-X)^T (\bar{X}-X)}{(X^T X)^2}$$

of the filter update vector decreases when a strong signal propagates across the array. If the updated filter vector were applied to the old data vector, moreover, the adaptive filter output would be

$$X^T A^{new} = X^T A^{old} \left[1 + \frac{2K_s X^T (\bar{X}-X)}{X^T X} \right]$$

$$= X^T A^{old} \left[1 - \frac{2K_s (\bar{X}-X)^T (\bar{X}-X)}{X^T X} \right] ,$$

so that the squared filter output would be reduced by a factor of $[1 - 2K_s (\bar{X}-X)^T(\bar{X}-X)/ X^TX]^2$ instead of $(1 - 2K_s)^2$. In the case of an ideal signal, there would be no attenuation at all. With weak signals having signal-to-noise ratios too low to trigger a filter freeze, this algorithm appears to have a definite advantage over the algorithm actually employed: the convergence factor K_s for the alternate algorithm would probably be significantly greater than for the algorithm actually used if signal degradation were the same for both algorithms, so that greater noise reduction and hence greater signal-to-noise gain might be achieved by the alternate algorithm. The suggested algorithm, like the one actually implemented, would also be less sensitive to power bursts from seismic events, so that non-repeating seismic events would be "forgotten" more rapidly than with the algorithm

$$A^{new} = A^{old} + 2\mu X^T A^{old}(\bar{X}-X) .$$

Normalization of the data vector X by its absolute value $|X|$ in the proposed algorithm instead of the absolute value $|\bar{X} - X|$, furthermore, is sensibly motivated: the adaptive filter responds to the phase characteristics of the data vector X and not to its magnitude. It is difficult to describe precisely the effect of the normalizing factor $|\bar{X} - X|$.

SECTION VI
VARIABILITY OF NOISE REDUCTION

A. INTRODUCTION

In Subsection B, the results of adaptive processing on two data samples are studied to determine the variation in noise reduction at different steer directions. In processing these samples, the adaptive filter is frozen when the similarity detection algorithm spots bursts of energy from the look direction. The reason for allowing the filter freeze is to simulate more accurately actual conditions of operation.

In Subsection C, the noise reduction figures from all noise samples processed for this report are compared in order to assess the range of signal-to-noise ratio improvement which may be expected from adaptive processing. An evaluation is made of the reasons for the performance achieved at each steer direction of each data sample. The signal-to-noise gain estimates for noise samples discussed in this subsection are only indirectly related to the super-directivity effects on coherent events demonstrated in Section V.

B. VARIATION IN NOISE REDUCTION AT DIFFERENT STEER DIRECTIONS

The first data sample covers the period 0115 to 0510 on day 321 of 1971. The noise level of this sample at periods between 15 and 20 seconds is close to the peak level for the 1971 autumn season. Sites 1, 2, 3, 6, 9, and 19 are used as input channels for the adaptive-filter beams, and the convergence rate is set to 0.5% of maximum. The adaptive-filter beams were aimed toward energy arriving at 3.5 km/sec from azimuth of 0°, 90°, 180°, and 270°, respectively. Since the adaptive filter set was frozen no fewer than seven times while the

steer direction was 90°, much of the energy in this four-hour sample must have come from azimuths close to 90°. Figure VI-1 presents two of the microseismic bursts which triggered cessation of the filter update. The period of oscillation is 18 seconds throughout both microseisms. If a detection algorithm is to distinguish between such microseismic activity and earthquakes, the dispersion characteristics of teleseismic events will have to be used. Broadband noise reduction for this noise sample is 1.2 dB at 0°, 1.4 dB at 90°, 3.1 dB at 180°, and 1.5 dB at 270°. Corresponding figures for noise reduction in the 15-to-43-second band are 1.7 dB, 1.6 dB, 4.1 dB, and 2.6 dB, respectively. Maximum noise reduction is 3.6 dB at the 14-second period for the 0° steer direction, 2.3 dB at a 20-second period for the 90° steer direction, 6.3 dB at an 18-second period for the 180° steer direction, and 4.9 dB at an 18-second period for the 270° steer direction. Figures VI-2 through VI-9 show beam spectra and noise reduction as a function of frequency, respectively, for the four look directions. Spectral levels are given in dB relative to one $(m\mu)^2/Hz$. The principal microseismic peak near 18 seconds is strongest on the 90° beam, but only slightly stronger than on the 180° beam. Beamsteer output levels on the 0° and 270° beams are a few dB lower at 18 seconds. A visual examination of the beam outputs reveals strong microseismic bursts of 18-second period on both the 90° beam and 180° beam. A 360° Fisher-detector scan of this data sample confirms that most of the energy is concentrated between 100° and 132° azimuths. The strong noise reduction relative to beamsteering on the 180° beam is due to the superdirectivity of the adaptive-filter beam. It is somewhat puzzling that noise reduction is noticeably better on the 270° beam than on the 0° beam. The WSW-ENE orientation of the six sites used as input channels may explain the difference: the main lobe of the array wavenumber response would be elongated along a NNW-SSE axis so that the 0° adaptive-filter beam would have more difficutly than the 270° adaptive-filter beam in rejecting noise from a 120° azimuth; the channel amplitude weightings would thus tend to be larger on the 0° beam than on the 270° beam, and

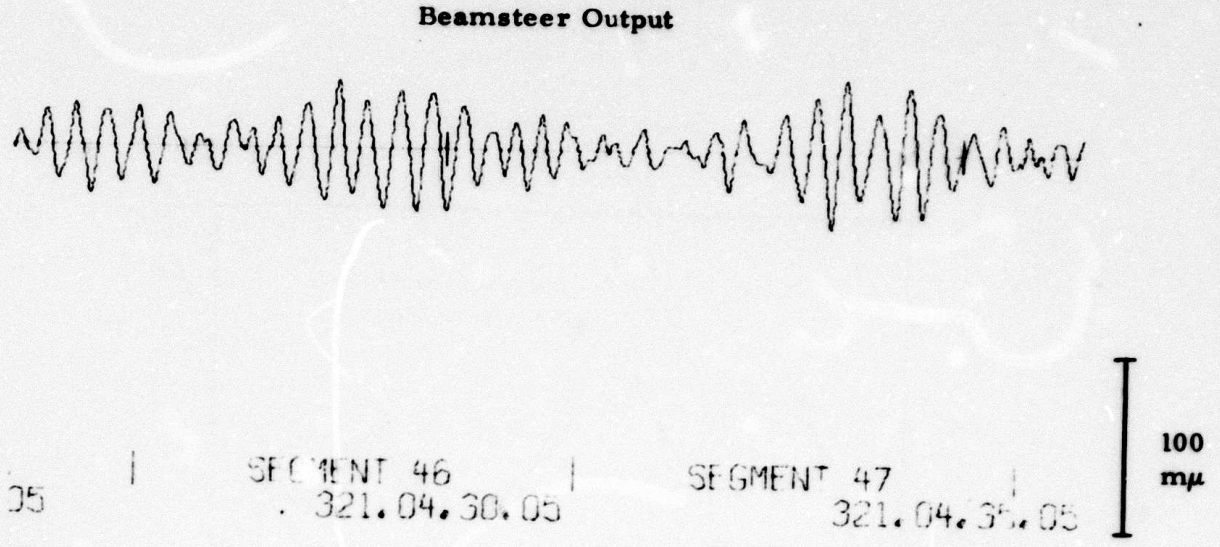


FIGURE VI-1

TWO MICROSEISMIC BURSTS WHICH TRIGGERED FILTER FREEZE
(DAY 321 1971, STEER DIRECTION 90°, $K_s = 0.005$)

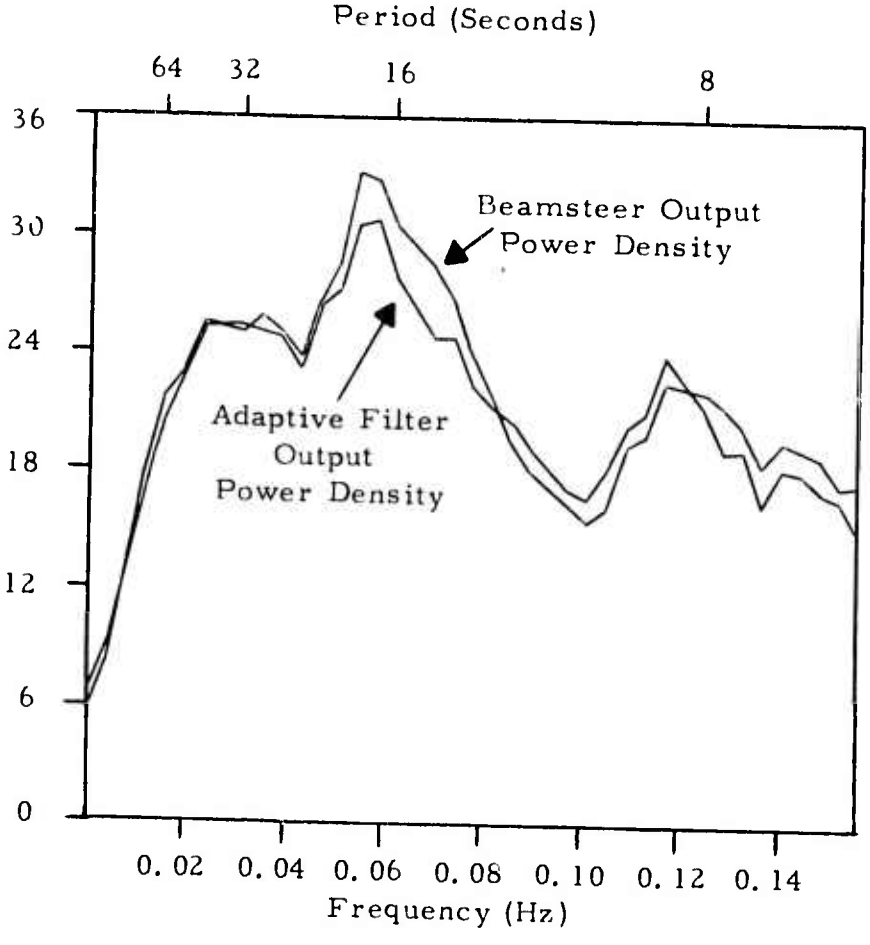


FIGURE VI-2

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 321 1971, STEER DIRECTION 0°, $K_s = 0.005$)

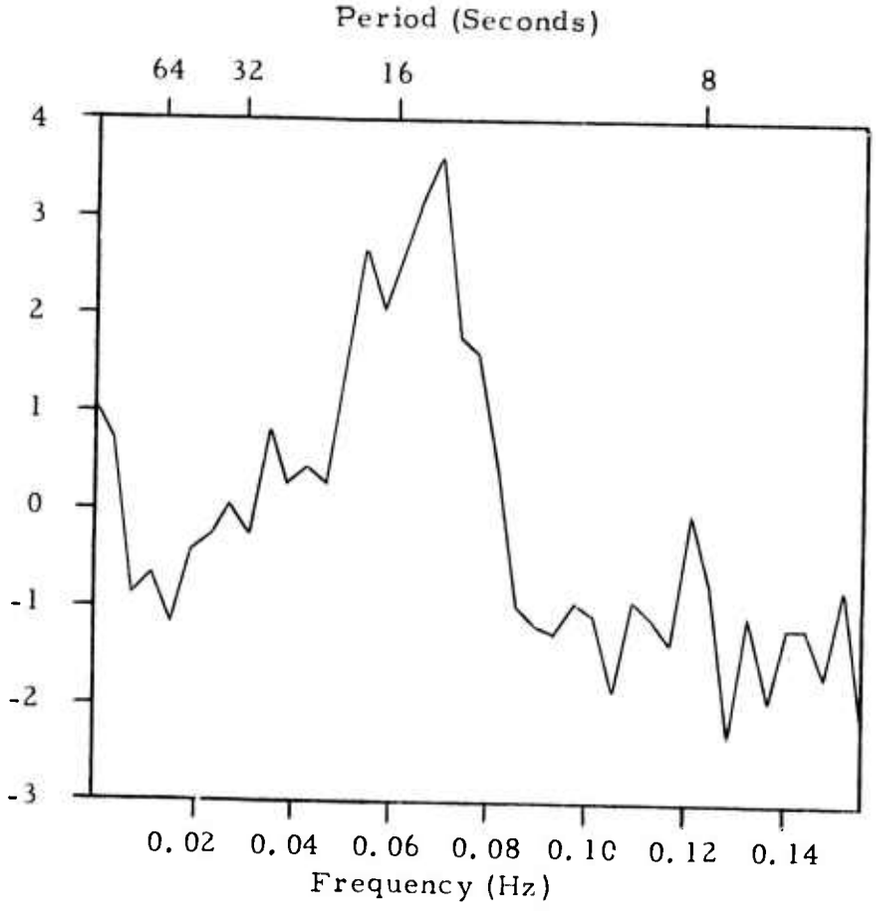


FIGURE VI-3

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY
(DAY 321 1971, STEER DIRECTION 0^{o}, $K_s = 0.005$)

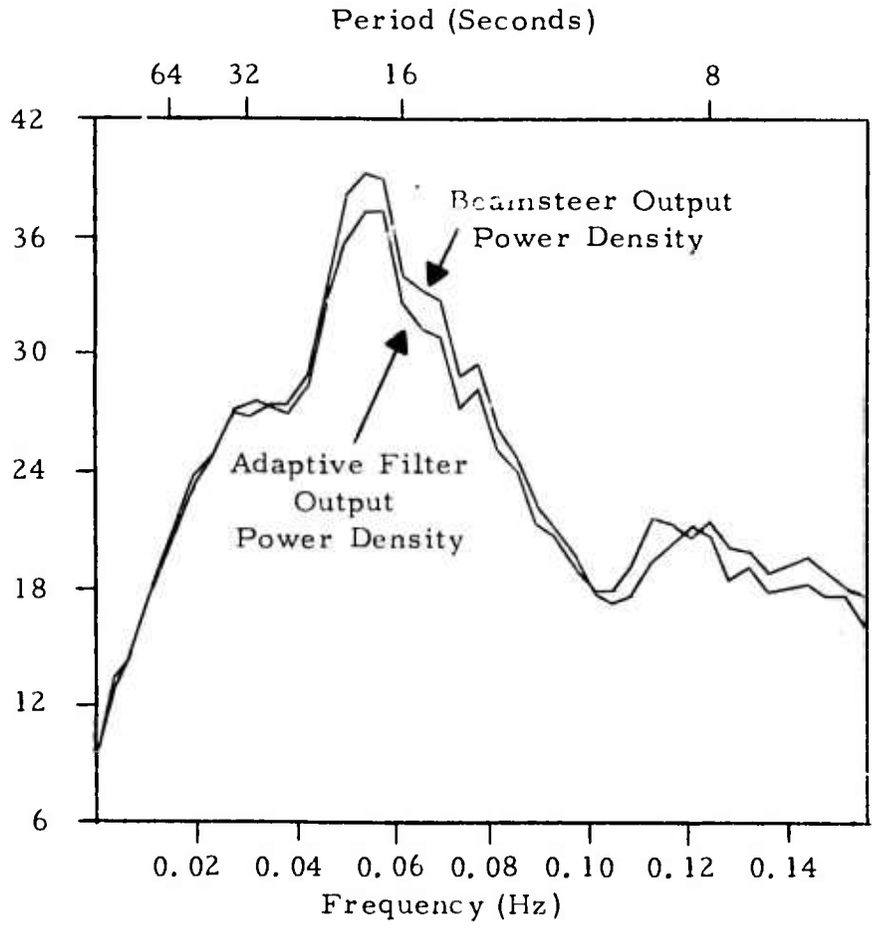


FIGURE VI-4

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 321 1971, STEER DIRECTION 90°, $K_s = 0.005$)

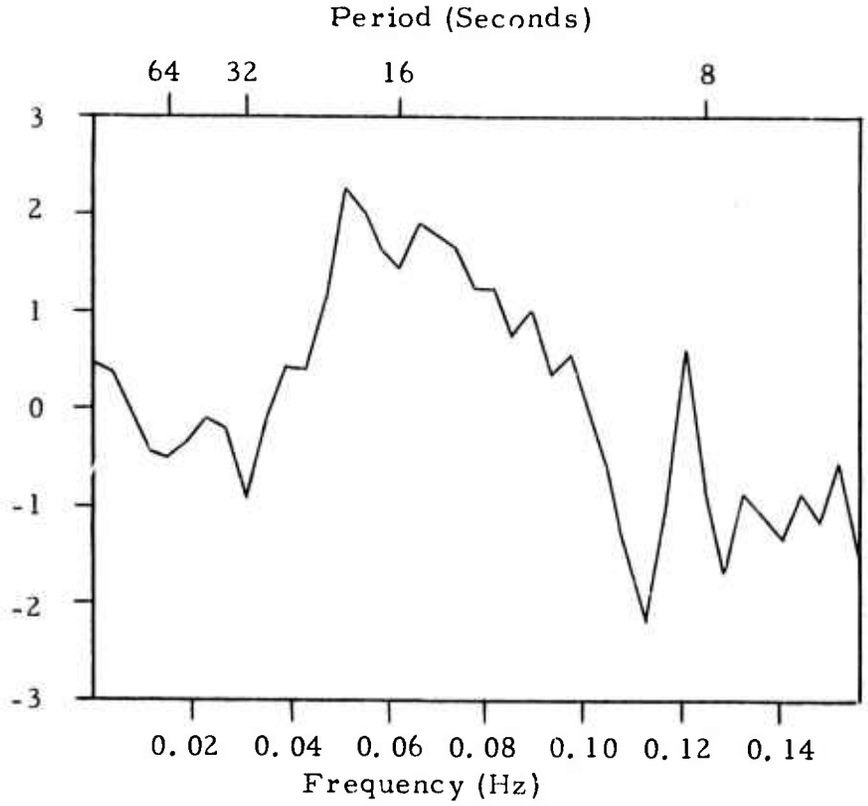


FIGURE VI-5

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY
(DAY 321 1971, STEER DIRECTION 90°, $K_s = 0.005$)

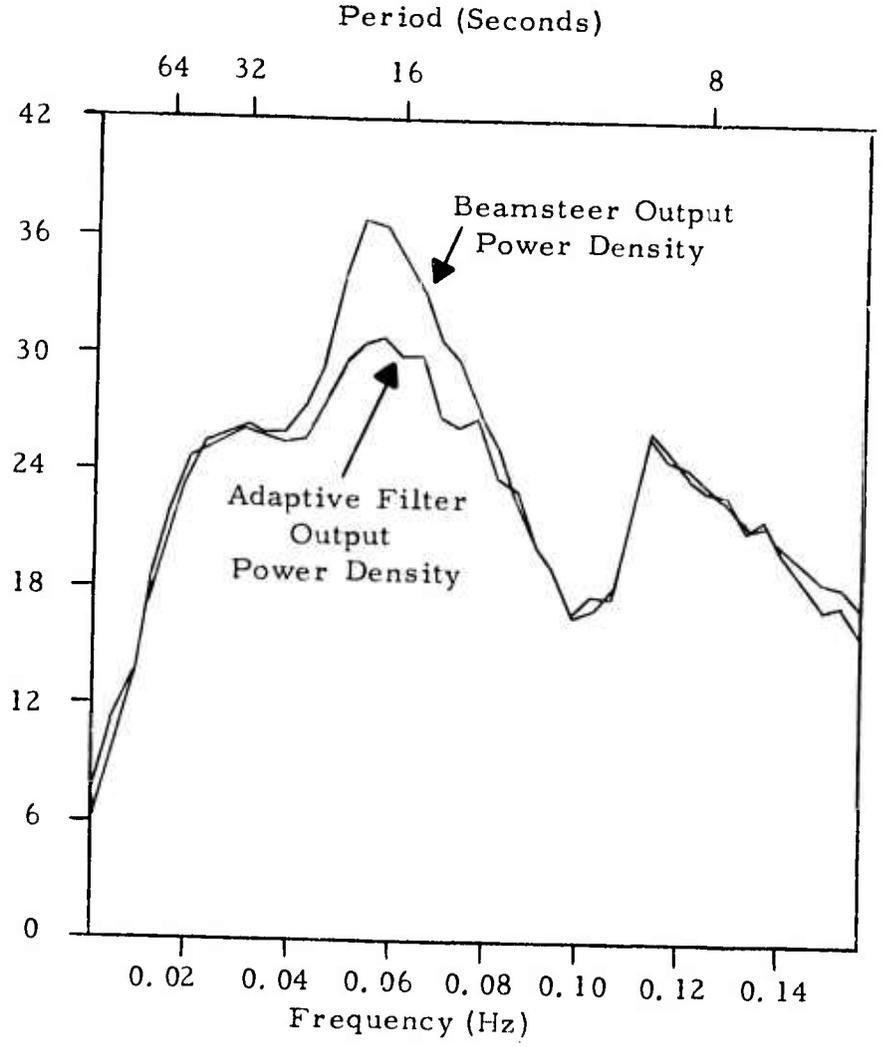


FIGURE VI-6

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 321 1971, STEER DIRECTION 180°, K_s = 0.005)

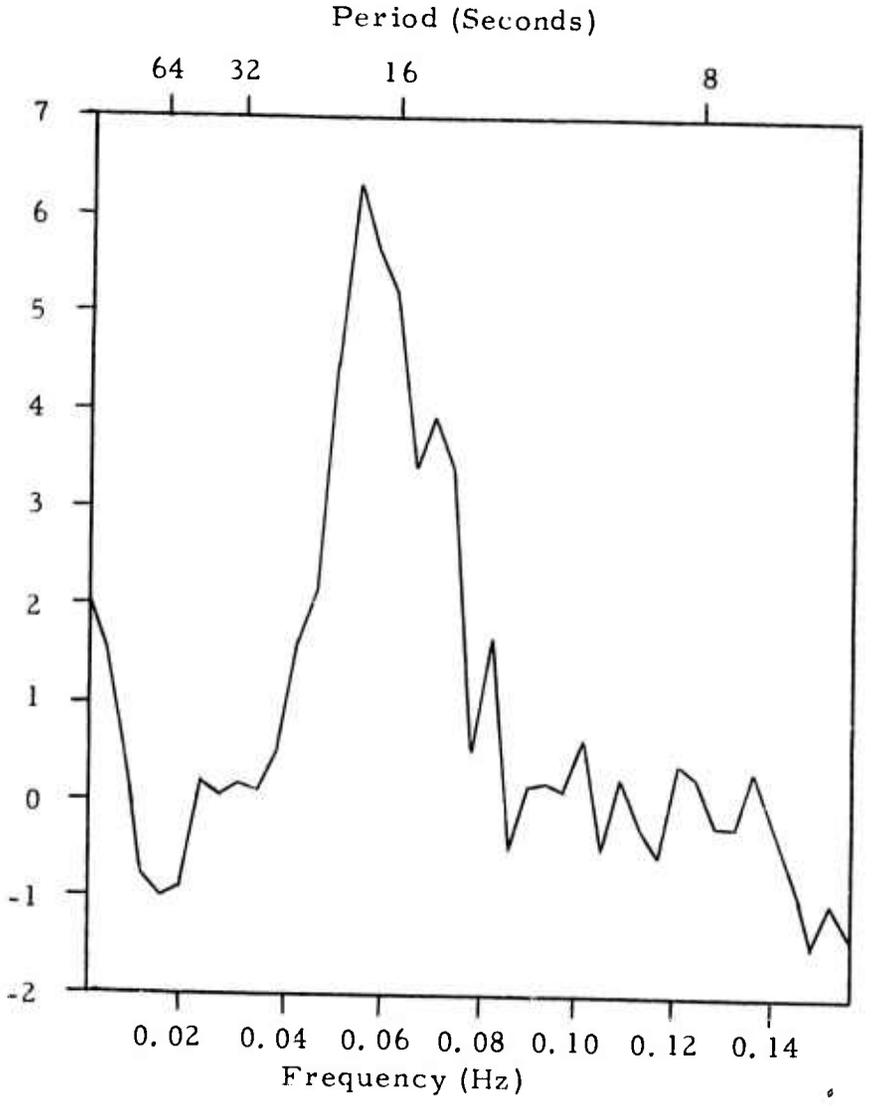


FIGURE VI-7

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY
(DAY 321 1971, STEER DIRECTION 180°, K_s = 0.005)

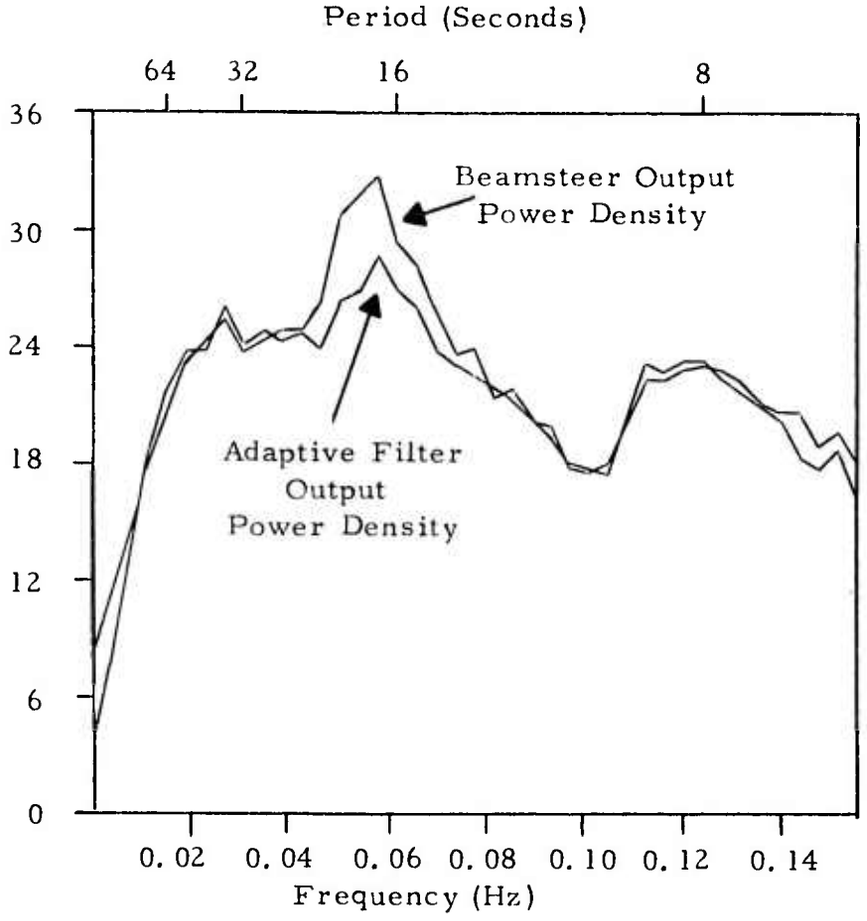


FIGURE VI-8

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 321 1971, STEER DIRECTION 270°, K_s = 0.005)

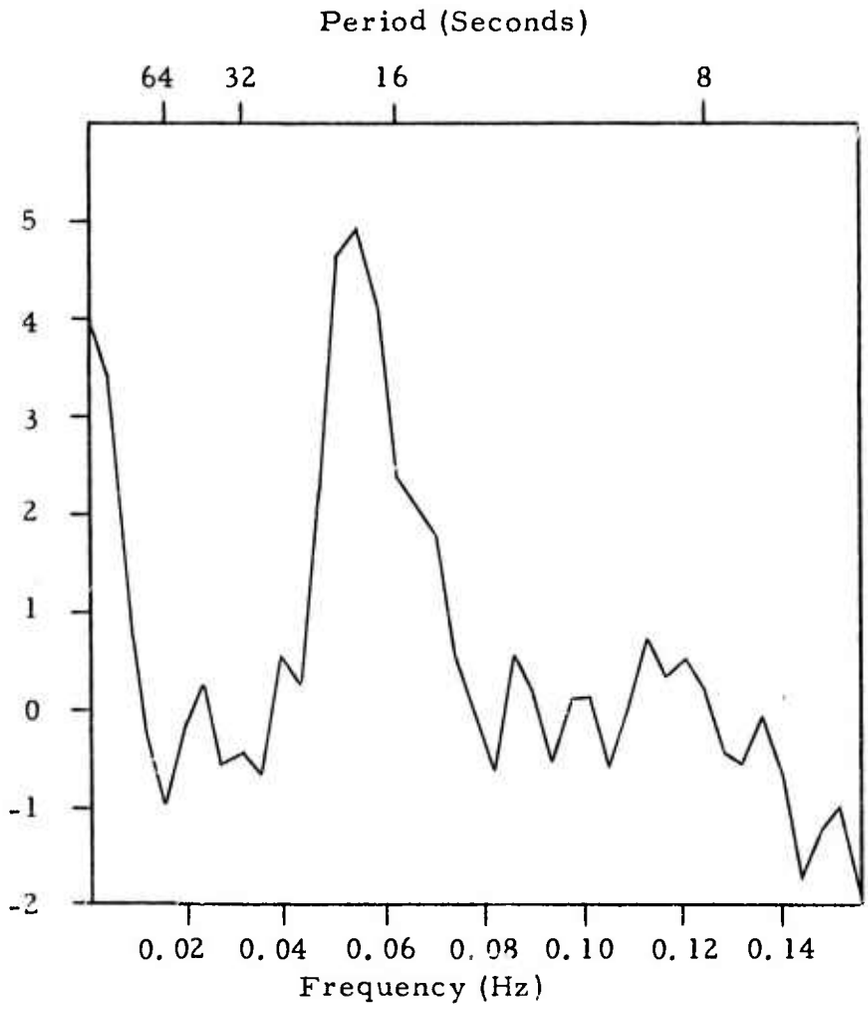


FIGURE VI-9

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY
(DAY 321 1971, STEER DIRECTION 270°, K_s = 0.005)

consequently poorer performance could be expected for noise of 8-second period, which tends to be relatively incoherent.

The second data sample spans the interval 0355 to 0725 on day 203 of 1971. This sample is the same sample described in Section IV (used there to study noise reduction as a function of convergence rate). Sites 8, 13, 14, 15, 16, and 17 of the ALPA array are the input channels for the beamsteer and adaptive-filter beams. The convergence rate is chosen to be 0.5% of maximum. Again, four beams are steered toward azimuths of 0°, 90°, 180°, and 270°. This noise sample contains at least four seismic events. Three of these are reported by the PDE bulletin and are given in Table VI-1. A fourth event which triggered a filter freeze on the 270° beam is displayed in Figure VI-10. This fourth event corresponds to a peak at 0426-0427 in the 275° look direction for a Fisher-detector scan from 180° to 300° at 5° azimuthal increments using 18 sites. The first New Ireland event arrived slightly less than an hour later and is pictured in Figures VI-11 and VI-12 as it appeared on the 270° and 180° beams. A glitch near the beginning of this event (seen most clearly in Figure VI-11) was caused by a burst of energy on a single channel. Although the event is located at an azimuth of 239° with respect to ALPA, the 180° beam contains the strongest signal. This fact suggests that the apparent direction of arrival is closer to 180° than 270°. A 360° Fisher-detector scan at 24° increments using sites 8, 13, 14, 15, 16, and 17 reaches its maximum at 228° before the glitch in Figures VI-11 and VI-12 and at 216° after the glitch. If such a shift in the apparent direction of arrival away from the direction of the source is a normal occurrence, the signal-to-noise ratio obtained by steering toward the source is lower than that which could be achieved by steering toward the apparent direction of arrival at ALPA. If the apparent direction of arrival were determined <u>before</u> steering the array, some processing gain might be realized. A Fisher-detector scan at 5° azimuthal increments from 180° to 300° using 18 sites, on the other hand, no longer exhibits the southward bias

TABLE VI-1

EVENTS ARRIVING AT ALPA BETWEEN
0355 AND 0725 ON JULY 22, 1971

Origin Time (Hr. -Min. -Sec.)	Lat (Deg.)	Long (Deg.)	Region	m_b	No. of Stations	Azimuth (Deg.)
04:37:09.1	5.7S	153.2E	New Ireland Region	—	8	239°
05:53:41.4	5.8S	153.6E	New Ireland Region	—	13	239°
06:07:52.8	0.4N	123.5E	Northern Celebes	5.4	18	270°

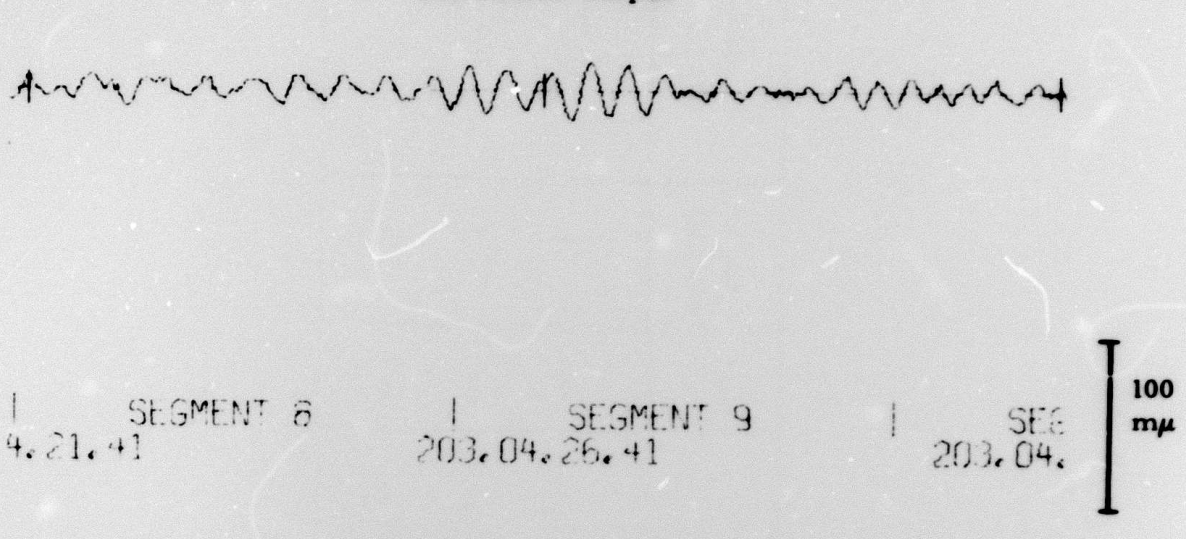


FIGURE VI-10

WEAK EVENT WHICH TRIGGERED FILTER FREEZE
(DAY 203 1971, STEER DIRECTION 270°, $K_s = 0.005$)

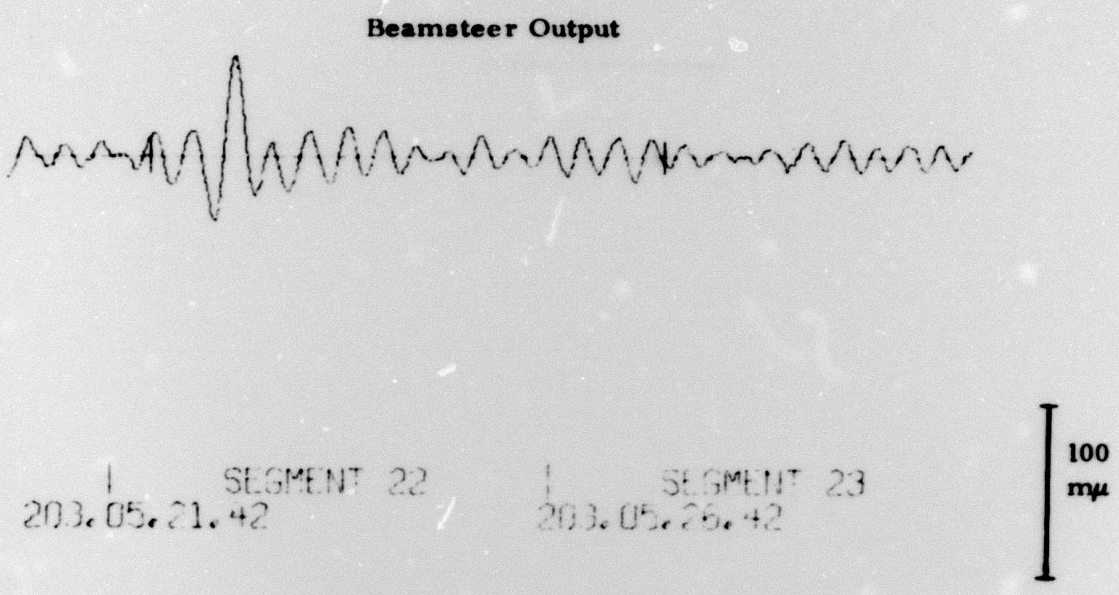


FIGURE VI-11

NEW IRELAND EVENT WITH GLITCH (ORIGIN TIME 04:37:09.1)
(DAY 203 1971, STEER DIRECTION 270°, $K_s = 0.005$)

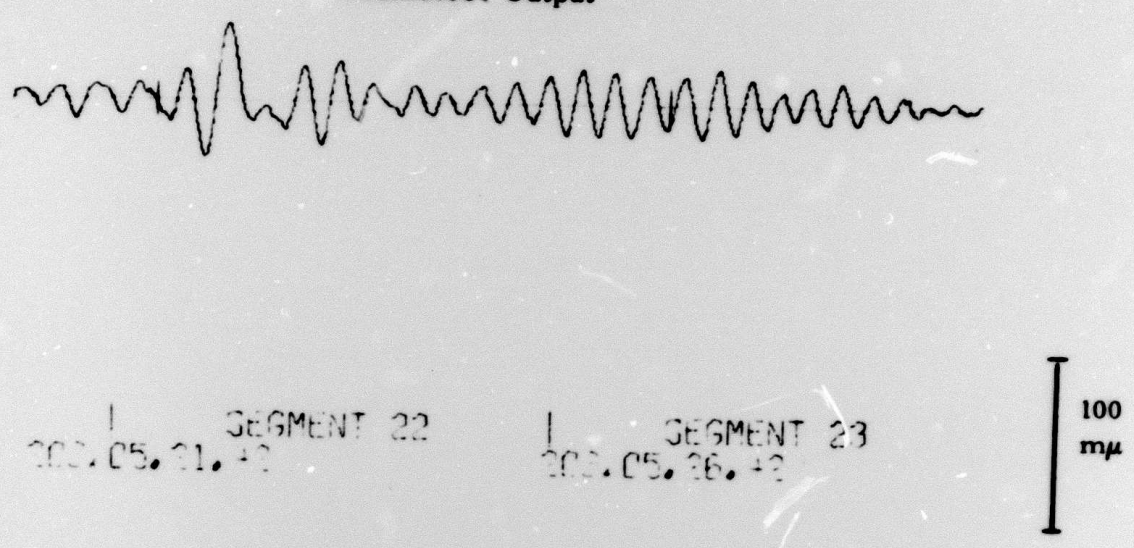


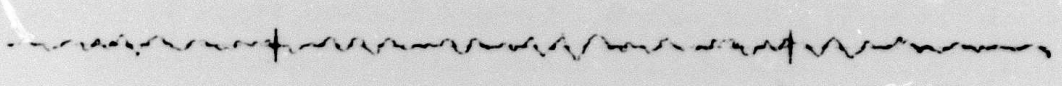
FIGURE VI-12

NEW IRELAND EVENT WITH GLITCH (ORIGIN TIME 04:37:09.1)
(DAY 203 1971, STEER DIRECTION 180°, $K_s = 0.0005$)

in the apparent arrival direction. The directional peak varies from 215° to 255° during the time interval 0517 to 0525. Thus, with the full array, the prospect of realizing signal-to-noise gain by steering toward the apparent direction of arrival seems limited. Figure VI-13 is a plot of the quietest noise section of the four-hour sample. A comparison with Figure VI-1 (from autumn) illustrates the wide variation in noise levels and spectral content throughout the year. A comparison of Figure VI-13 with the seismic events during the period 0355 to 0725, on the other hand, indicates that a significant portion of the total energy in the sample is from the seismic events. Since PDE events occur slightly more frequently than once every two hours, this situation may be typical of summertime. Figure VI-14 portrays the second New Ireland event as seen by the 270° beam. On the 180° beam, the energy level is about the same as on the 270° beam. The Fisher-detector sweep using sites 8, 13, 14, 15, 16, and 17 peaks at a 228° azimuth, but at a 237.5° azimuth when 18 sites are used. Figure VI-15 is the magnitude 5.4 event from the Northern Celebes region as seen by the 270° beam. Once again, the six-site Fisher-detector scan indicates an arrival azimuth of 252°, while the corresponding 18-site scan produces azimuthal estimates ranging from 250° to 285° over the time interval 0655-0700.

Broadband noise reduction for this sample is 2.2 dB at 0°, 1.4 dB at 90°, 2.4 dB at 180°, and 2.4 dB at 270°. The corresponding noise reduction in the 15-to-43-second band is 2.4 dB, 1.4 dB, 2.5 dB, and 2.6 dB, respectively. Maximum noise reduction is 3.7 dB at a 23-second period for the 0° beam, 3.3 dB at a 51-second period for the 90° beam, 5.5 dB at a 23-second period for the 180° beam, and 4.2 dB at a 21-second period for the 270° beam. Figures VI-16 through VI-21 graph beam spectra and noise reduction as a function of frequency, respectively, for the 0°, 90°, and 180° beams. Since noise reduction on the 270° beam seldom differs by more than 0.1 dB from that shown in Figure IV-6, no new illustrations have been generated for that

Beamsteer Output



SEGMENT 32
203.06.06.42

SEGMENT 33
203.06.11.42

SEG

100 mμ

Adaptive Filter Output

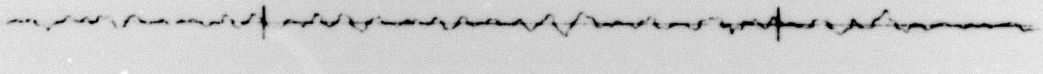
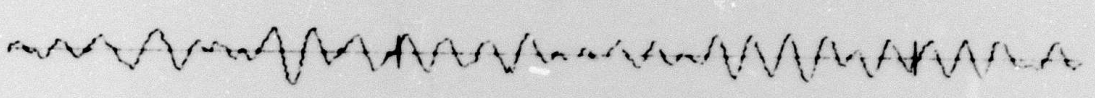
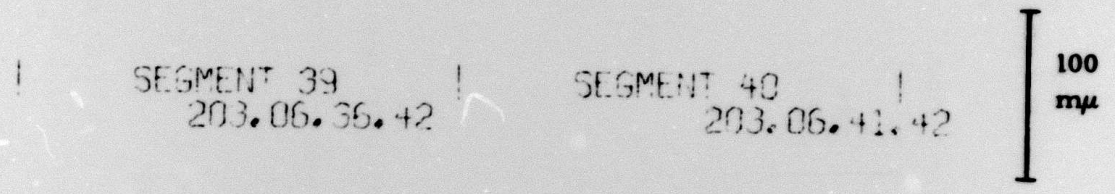


FIGURE VI-13

QUIESTEST NOISE SECTION BETWEEN 0355 and 0725 (DAY 203 1971, STEER DIRECTION 270°, $K_s = 0.005$)

Beamsteer Output





Adaptive Filter Output

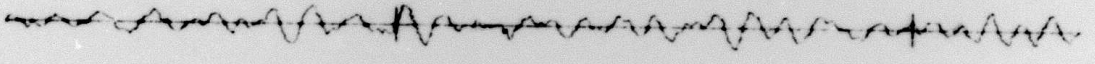


FIGURE VI-14

NEW IRELAND EVENT (ORIGIN TIME 05:53:41.4)
(DAY 203 1971, STEER DIRECTION 270°, $K_s = 0.005$)

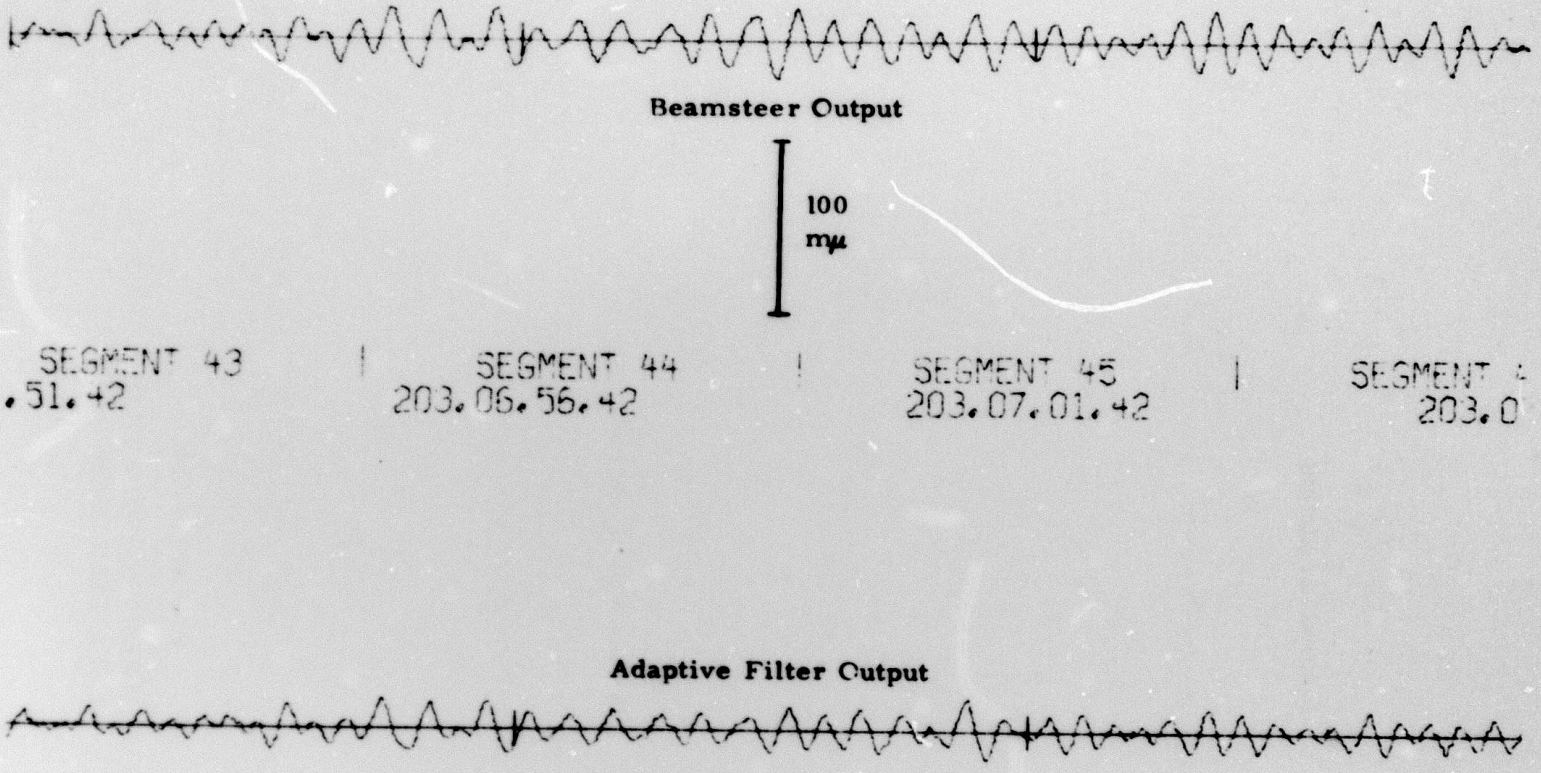


FIGURE VI-15

MAGNITUDE 5.4 NORTHERN CELEBES EVENT
(DAY 203 1971, STEER DIRECTION 270°, $K_s = 0.005$)

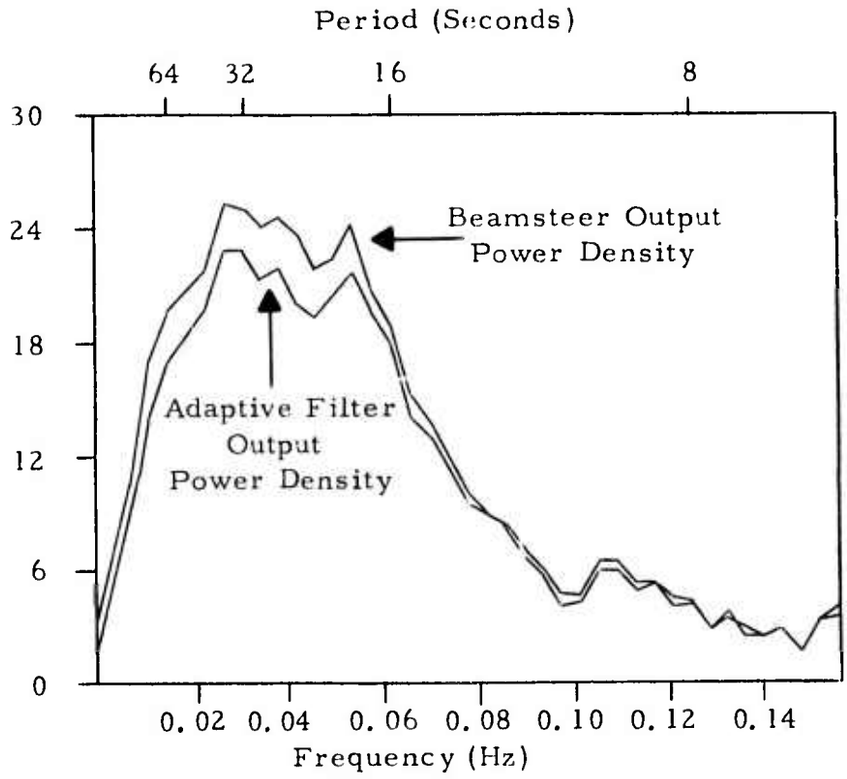


FIGURE VI-16

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 203 1971, STEER DIRECTION 0^{o}, $K_s = 0.005$)

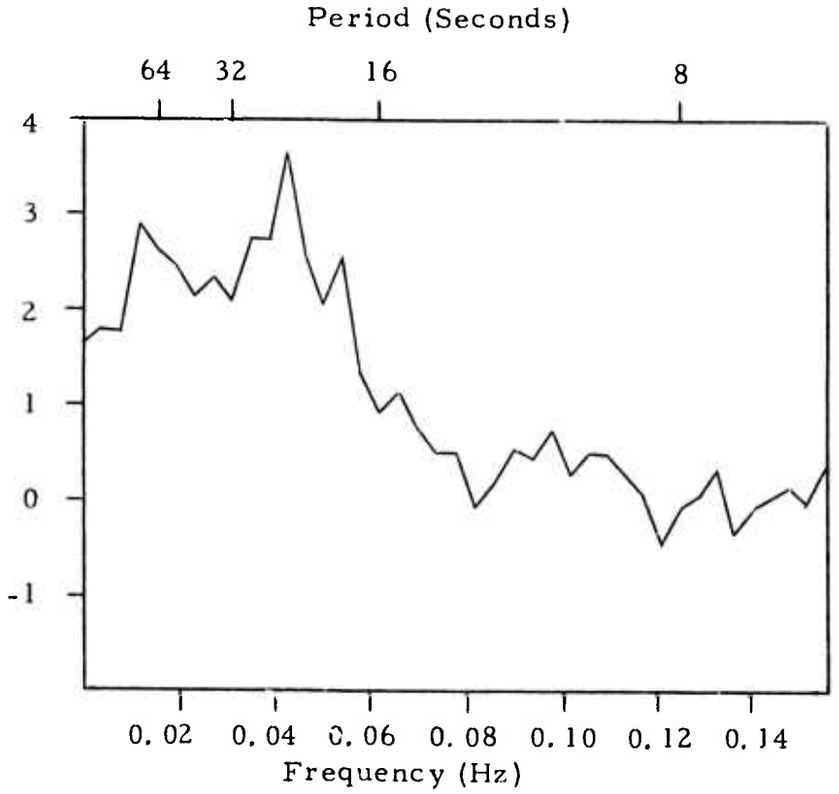


FIGURE VI-17

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY
(DAY 203 1971, STEER DIRECTION 0^{o}, $K_s = 0.005$)

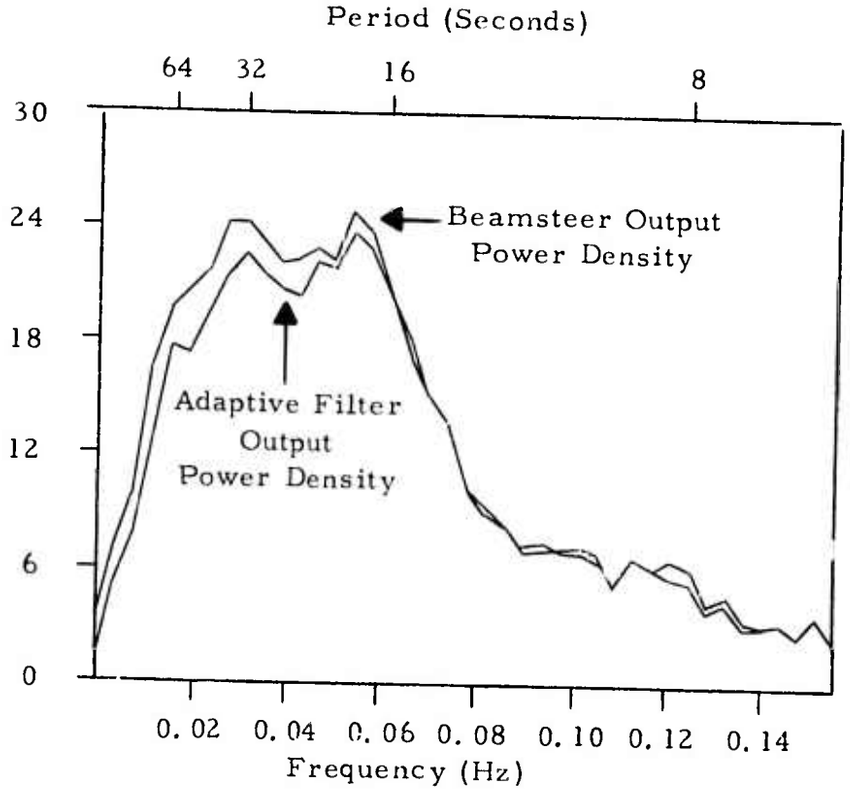


FIGURE VI-18

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 203 1971, STEER DIRECTION 90°, K_s = 0.005)

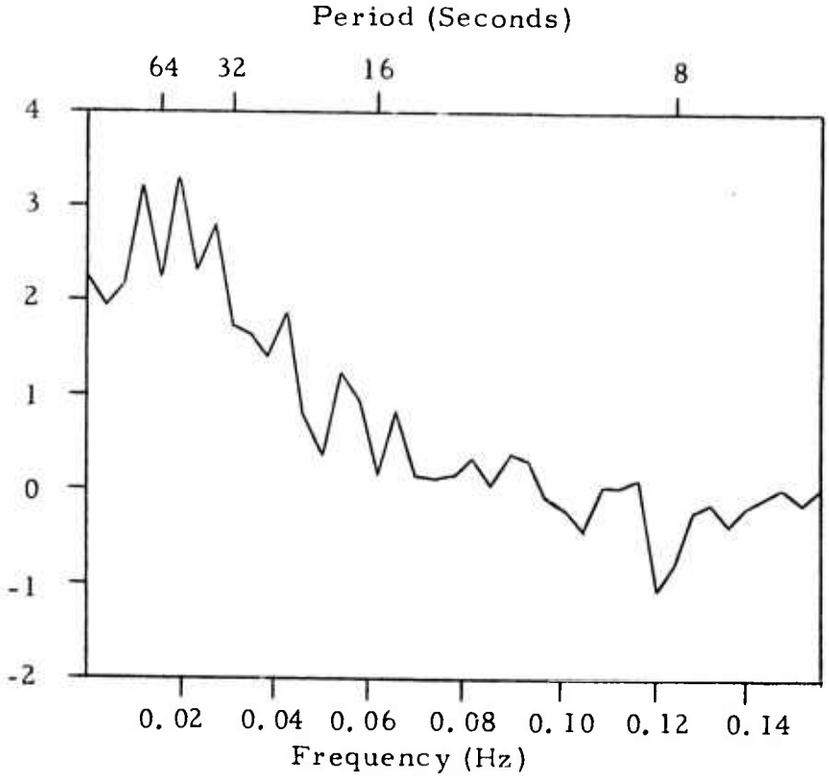


FIGURE VI-19

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY
(DAY 203 1971, STEER DIRECTION 90^{o}, $K_s = 0.005$)

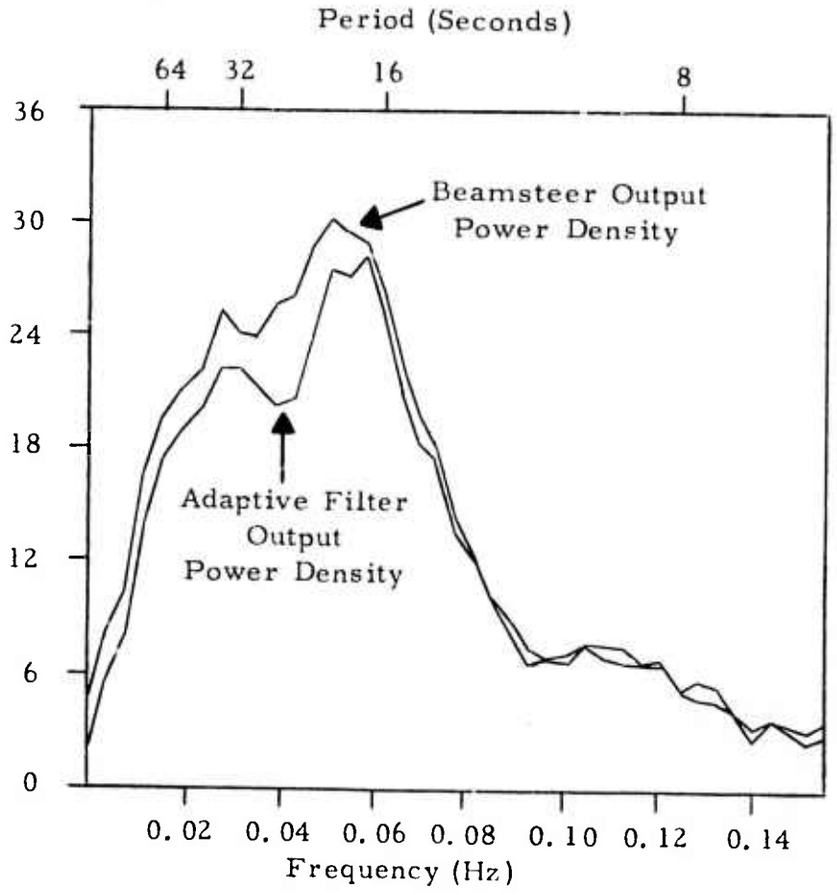


FIGURE VI-20

BEAMSTEER AND ADAPTIVE FILTER OUTPUT POWER DENSITY
(DAY 203 1971, STEER DIRECTION 180°, K_s = 0.005)

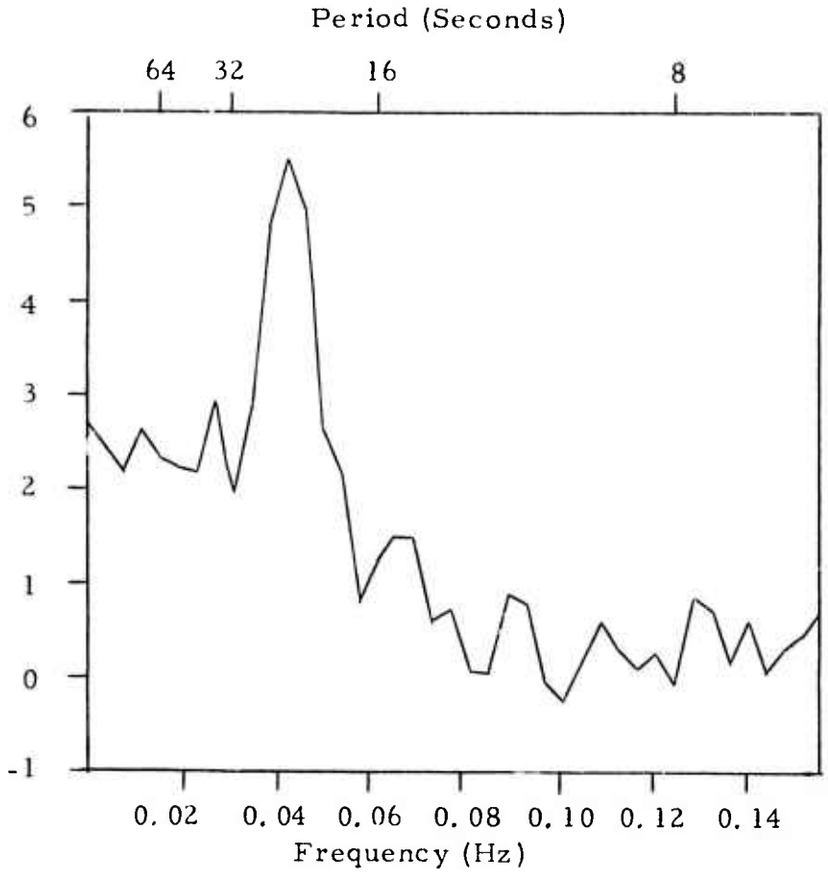


FIGURE VI-21

ADAPTIVE FILTER NOISE REDUCTION AS A FUNCTION OF FREQUENCY
(DAY 203 1971, STEER DIRECTION 180°, K_s = 0.005)

beam. The beams aimed toward 0^{o} and 90^{o} possess the least power: spectral peaks in the 20-to-36 second band are comparable to the height of the principal microseismic peak at 18 seconds. The main peaks at 20 seconds on the 130^{o} beam and at 21 seconds on the 270^{o} beam are about 5 dB higher. Note that the 180^{o} beamsteer output contains less power at periods of 21 seconds and above tnan the 270^{o} beam. The relatively high-frequency power on the 180^{o} beam probably comes from the two New Ireland events. Since most of the energy is this sample arrives at apparent azimuth ranging from 210^{o} to 270^{o}, the highest noise reduction peaks for the 180^{o} and 270^{o} beams may be attributed to superdirectivity effects on coherent seismic events. The greater noise reduction on the 0^{o} beam is due to the fact that the 90^{o} beamsteer output contains less coherent energy to eliminate.

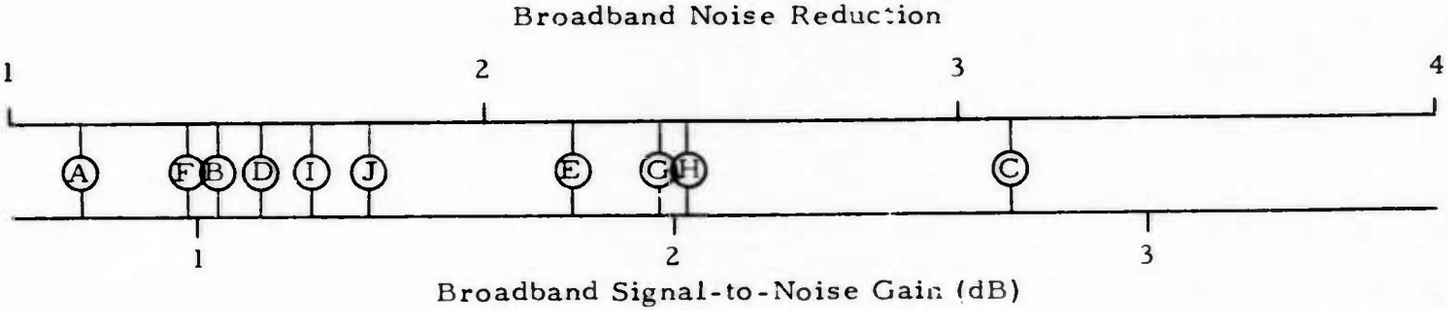
C. VARIABILITY OF SIGNAL-TO-NOISE GAIN

Table VI-2 furnishes the noise reduction values for each of the ten four-hour samples processed for this report. Figure VI-22 provides a plot of these noise reduction values together with the corresponding signal-to-noise gains for the weak signal from 300^{o} - 305^{o} (day 276), which is 6 dB above noise level on the beamsteer output (see Subsection IV-D). Sample I suggests that 1.2 dB broadband and 1.3 dB narrowband signal-to-noise gain can be expected from pure summer background noise. Sample J intimates that the corresponding gains for winter background noise are 1.4 dB and 1.7 dB, respectively. Cases A through D near the peak fall noise level of 1971 on day 321 illustrate the variation in noise reduction as steer direction changes. Superdirectivity effects account for the sustained 2.7 dB and 3.7 dB gains on the 180^{o} beam. Since the filter set was frozen seven times on the 90^{o} beam, much of the energy is very close to the look direction and adaptive processing produces signal-to-noise gains of only 1.0 dB broadband and 1.2 dB over the 15-to-43-second band. On the 270^{o} beam, the gains achieved are 1.1 dB and 2.2 dB. In this case,

TABLE VI-2

NOISE REDUCTION FOR TEN FOUR-HOUR SAMPLES

Sample	Day	Interval	Steer Direction	Noise Reduction (dB)	
				Broadband	Narrowband
A	321/1971	0115-0510	0°	1.156	1.702
B			90°	1.442	1.608
C			180°	3.111	4.136
D			270°	1.535	2.584
E	203/1971	0355-0725	0°	2.192	2.374
F			90°	1.380	1.392
G			180°	2.370	2.518
H			270°	2.423	2.576
I	238/1970	0757-1150	270°	1.639	1.746
J	007/1972	0238-0616	270°	1.763	2.137



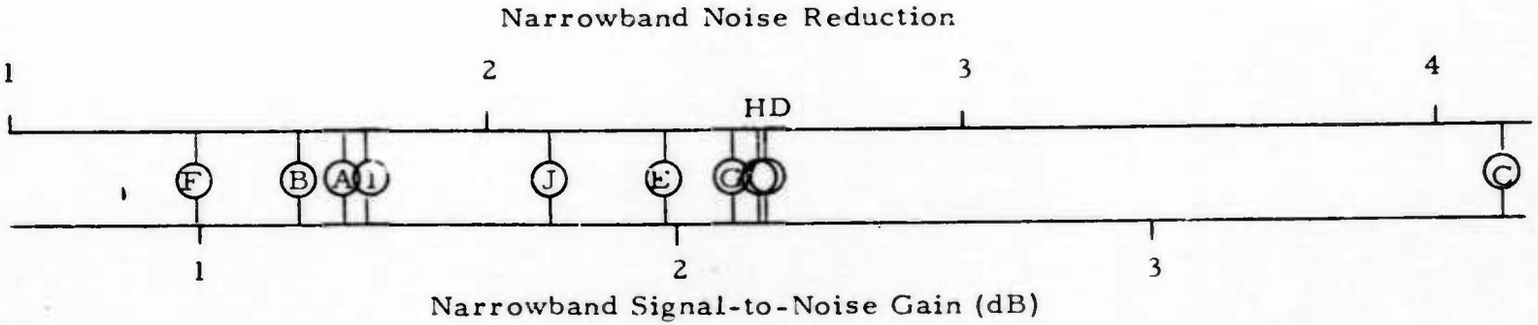


FIGURE VI-22

NOISE REDUCTION AND SIGNAL-TO-NOISE GAIN (6 dB S/N RATIO) FOR EACH OF THE TEN FOUR-HOUR NOISE SAMPLES PROCESSED

adaptive filtering was able to shape the array beam pattern so as to reduce significantly the microseismic energy from 108° to 132°. The 0° adaptive-filter beam was not nearly so successful and yielded gains of 0.8 dB and 1.3 dB. A possible explanation is the orientation of the main lobe of the array wavenumber response along a NNW-SSE axis.

Cases E through H demonstrate the signal-to-noise gain achievable at the steer directions 0°, 90°, 180°, and 270° when several events propagate across ALPA at apparent azimuths ranging from 210° to 270°. Superdirectivity accounts for the 2.0 dB and 2.1 dB gains on the 180° beam as well as the 2.0 dB and 2.2 dB gains on the 270° beam. Some of the 1.8 dB and 2.0 dB gain on the 0° beam can be explained in the same way. In contrast, signal-to-noise gains of only 1.0 dB (both broadband and narrowband) were obtained on the 90° beam.

These results imply that signal-to-noise ratio improvement of one to two dB can be preserved over four-hour noise samples. In one case involving superdirectivity, broadband gain was 2.7 dB, narrowband gain 3.7 dB. Of course, as was amply demonstrated in Section V, even greater signal-to-noise improvement can be achieved on off-azimuth seismic events of brief duration.

SECTION VII
CONCLUSIONS

Floating DC levels in the data channels transmitted from ALPA caused considerable difficulty in implementing an adaptive filtering system until their effect was studied theoretically and effective remedial action taken. Two steps were necessary:

- The data traces were run through a filter having a response exactly equal to zero at DC.
- The adaptive filtering program was examined to uncover DC bias introduced by the computations. Bias compensation was incorporated into the program and intermediate results were rounded instead of truncated whenever possible.

Roundoff error in the adaptive-filter update equation

$$A^{new} = A^{old} + \frac{2K_s X^T A^{old} (\bar{X} - X)}{(\bar{X} - X)^T (\bar{X} - X)}$$

almost became a serious problem in obtaining the results of this report. When the data points were scaled by a factor of 16, however, error in the input channels to the adaptive filter was reduced to the point where the data vector X and the beamsteer output vector $\bar{X}$ were almost as accurate as the corresponding vectors computed with floating-point arithmetic. In one noise sample, this

scaling procedure increased the noise reduction of adaptive filtering relative to beamsteering from 2 dB to 6 dB at the convergence rate $K_s = 0.30$. The 2-dB noise reduction figure for unscaled data is apparently due to sign reversals in the adaptive filter output $y(t) = X^T A$ (so that the adaptive filter vector A moves in the wrong direction) and to zero values y(t) after roundoff (so that the adaptive filter vector does not move at all). With data scaled by 16, errors in the vector $(\overline{X} - X)$ and the adaptive filter output y(t) were predominantly digitization errors. The sensor gain doubling at ALPA in the summer of 1972 should permit these errors to be halved. At the most frequently employed convergence rates (near $K_s = 0.005$), the dominant source of error was the error in rounding the updated filter weights (on the right side of the update equation) to the nearest filter-weight count. The mean angle of error in the vector $(A^{new} - A^{old})$ was estimated as 18^o for one summer noise sample at the convergence rate $K_s = 0.005$. This angle could have been reduced to 10^o if errors in the maximum-likelihood constraint conditions had been corrected differently. As the convergence rate drops below 0.5%, the filter-weight roundoff error becomes progressively worse until ultimately the filter vector A cannot change. The way to improve this situation is to incorporate more bits into the filter-weight representation. Such a solution would have meant abandoning the special convolution-filter microcode instruction incorporated in the IBM 360/40 computers at SDAC, where a 16-bit filter-weight representation is required for the CFIL microcode. Had filter-weight roundoff error been eliminated, the effect would have been to reduce (probably only slightly) the convergence rate at which the highest adaptive-filtering signal-to-noise gains relative to beamsteering were achieved.

In determining adaptive-filtering signal-to-noise gains, the critical area of concern is the processing improvement for weak signals in the borderline detection range, where the signal-to-noise ratio on the beamsteer output is between 6 and 12 dB. With such weak signals, detection procedures cannot consistently recognize the presence of a signal, and no filter-freeze procedure can be implemented. One signal approximately 6 dB above

the noise level on the beamsteer output was used for the critical results. Signal-to-noise gain was measured as the difference between adaptive-filtering noise reduction and signal degradation. Optimum gain was realized near a 0.5% convergence rate. Using noise data from day 238 of 1970, signal-to-noise gain for the weak signal was 1.23 dB broadband, 1.34 dB in the frequency band for periods between 43 and 15 seconds. With noise data from day 203 of 1971, broadband gain was 2.03 dB, narrowband gain 2.19 dB. Due to greater degradation of stronger signals, the signal-to-noise gain was lower for signals 18 to 24 dB above the noise level on the beamsteer output when the adaptive filter set was permitted to update.

With signals as strong as these, it is easy to detect their presence. A scaled version of the Fisher detection algorithm was used for this purpose. With the particular adaptive algorithm employed, the standard procedure is to freeze the adaptive filter set. When the filter set was prevented from updating upon signal detection, signal degradation was less than 0.1 dB for a signal 18 dB above the noise level on the beamsteer output and almost exactly 0 dB for a signal with a beamsteer-output signal-to-noise ratio of 24 dB. These figures are contrasted with a signal degradation of 0.41 dB for the 6-dB signal at a convergence rate of 0.5%. Although the signal degradation is lower for the two strong signals when the filter is frozen, noise reduction begins to drop as the elapsed time from the point of the filter freeze increases. Loss in noise reduction was measured by twice processing a noise sample from day 232 of 1970. In both cases, the filter was permitted to adapt for the first three hours of the noise sample. In the first computer run, it was allowed to adapt for one more hour. In the second run, it was frozen during the final hour. The apparent trend of the loss in noise reduction indicates that higher signal-to-noise gain is preserved for at least 15 minutes by freezing the filter rather than updating it in the case of the 18-dB signal. Superior gain is maintained much longer for the 24-dB signal.

The alternate adaptive algorithm

$$A^{new} = A^{old} + \frac{2K_s X^T A^{old} (\bar{X} - X)}{X^T X}$$

adapts much less rapidly in the presence of a signal. It is very possible that this algorithm could operate at convergence rates higher than the value $K_s = 0.005$ without degrading signals any more than the implemented algorithm does at its optimum rate of $K_s = 0.005$. If so, greater signal-to-noise gains could be realized through increased noise reduction at higher convergence rates.

One data sample from 2000 to 2357 on day 276 of 1971 was especially rich in signals (both on-azimuth and off-azimuth). The following conclusions can be drawn from this four-hour data sample:

- Off-azimuth events are strongly suppressed in the adaptive-filter beam when the filter set is not frozen. Some off-azimuth signals are virtually annihilated. The more powerful the off-azimuth event, the more it is stifled.
- Greater directional resolution at ALPA and other similar long-period arrays is achievable through the beam-narrowing capability of multichannel filtering. This fact is extremely important if long-period arrays of this type are to be used for signal detection and location or for separation of multiple events. With time-varying adaptive filters, off-azimuth events can be nulled out in proportion to their signal-to-noise ratio with possible complications if two events overlap in time. With fixed non-varying multichannel filters, superdirectivity can be preserved in all circumstances, but then the ability to quell specific bursts of off-azimuth energy in an on-line processing mode is impaired.

Four four-hour noise samples were processed for this report. Adaptive filter beams were steered toward four different look directions in processing two of these noise samples. In eight out of ten cases, the broadband signal-to-noise gain which would have achieved for the weak 6-dB signal was within the range 0.98 to 2.02 dB. In one case, it would have been 0.75 dB; in another, it would have been 2.7 dB. Over the band 0.0234 to 0.0664 Hz (corresponding to periods between 43 and 15 seconds), the signal-to-noise gain for the weak signal would have been between 0.98 and 2.0 dB in six cases, between 2.0 dB and 2.2 dB in three cases. The last case would have yielded a narrowband gain of 3.74 dB. The narrowband values are meaningful if a bandpass filter for periods between 40 and 15 seconds is applied to the data.

SECTION VIII
REFERENCES

Blandford, R. R., 1970, "An Automatic Event Detector at TFO", Seismic Data Laboratory Report No. 263, Teledyne Geotech.

Blandford, R. R., 1972, "Qualitative Properties of the F-Detector", Seismic Data Laboratory Report No. 291, Teledyne Geotech.

Booker, A. H., 1965, "Analysis of Variance as a Method for Seismic Signal Detection", Seismic Data Laboratory Report No. 216, Teledyne UED.

Daniell, T. P., 1968, "Adaptive Estimation with Mutually Correlated Training Samples", Stanford Electronic Laboratories, Stanford, California, Document SEL-68-083 (Technical Report TR 6778-4).

Edwards, J. P. III, S. A. Benno, and G. Creasey, 1967, "Evaluation of the CPO Auxiliary Processor", CPO Special Report No. 5, Texas Instruments Incorporated.

Frost, O. L. III, 1972, "An Algorithm for Linearly Constrained Adaptive Array Processing ", Proceedings of the IEEE, Volume 60, pp. 926-935.

Gnedenko, B. V., 1962, The Theory of Probability, Chelsea Publishing Company, New York, pp. 132-135.

Melton, B. S., and L. F. Bailey, 1957, "Multiple Signal Correlators", Geophysics, Volume 22, No. 3, pp. 565-588.

Shumway, R. H., and L. L. Husted, 1970, "Frequency-Dependent Estimation and Detection for Seismic Arrays", Seismic Data Laboratory Report No. 249, Teledyne Geotech.

Shumway, R. H., 1971, "On Detecting a Signal in N Stationarily Correlated Noise Series", Technometrics, Volume 13, pp. 499-520.

Shumway, R. H., 1972, "Some Applications of a Mixed Signal Processor", Seismic Data Laboratory Report No. 280, Teledyne Geotech.

Smart, E., and E. A. Flinn, 1971, "Fast Frequency-Wavenumber Analysis and Fisher Signal Detection in Real-Time Infrasonic Array Data Processing", Geophysical Journal of the Royal Astronomical Society, Volume 26, pp. 279-284.

Smart, E., 1972, "FKCOMB, A Fast General-Purpose Array Processor", Seismic Array Analysis Center Report No. 9, Teledyne Geotech.

Widrow, B., 1966, "Adaptive Filters I: Fundamentals", Stanford Electronics Laboratories, Stanford, California, Report SEL-66-126 (Technical Report 6764-6).

Wirth, M. H., R. R. Blandford, and R. H. Shumway, 1971, "Automatic Network Detection", Seismic Data Laboratory Report No. 285, Teledyne Geotech.

Wirth, M. H., 1971, "Computation of Multiple-Event Probabilities", Seismic Data Laboratory Report No. 277, Teledyne Geotech.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD 785 199

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SIMULATED ON-LINE ADAPTIVE PROCESSING RESULTS USING ALASKA LONG PERIOD ARRAY DATA		5. TYPE OF REPORT & PERIOD COVERED Special 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Thomas E. Barnard		8. CONTRACT OR GRANT NUMBER(s) F33657-72-C-0725
9. PERFORMING ORGANIZATION NAME AND ADDRESS Texas Instruments Incorporated Equipment Group Dallas, Texas 75222		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS VELA T/2705/B/ASD
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency Nuclear Monitoring Research Office Arlington, Virginia 22209		12. REPORT DATE 23 October 1973 13. NUMBER OF PAGES ~~298~~ 299
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Air Force Technical Applications Center VELA Seismological Center Alexandria, Virginia 22314		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

ARPA Order No. 1714

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Adaptive multichannel filtering
Seismic signal processing
Roundoff error
Alaska Long-Period Array

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report deals with results obtained from operating an adaptive time-domain maximum-likelihood filtering system on data from the Alaska Long-Period Array (ALPA). Signal-to-noise gain of adaptive filtering relative to beamsteering is investigated as a function of convergence rate and steer direction. In addition, the effect upon signal-to-noise gain of freezing the adaptive filter set is described. Both on-azimuth and off-azimuth signals are examined to determine how much they are attenuated in the adaptive

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

1.

20. continued

filter beam in comparison with the beamsteer output. Signal-to-noise gain values presented are measured using beam output traces formed from actual seismic data. Theoretical studies of the effect of floating DC levels and roundoff error are also contained in this report.

A major advantage of adaptive multichannel filtering is found to be the capability to narrow the main lobe of the array beam pattern.